



MÁSTER EN LETRAS DIGITALES

TRABAJO DE FIN DE MÁSTER

CURSO 2017/2018

DISEÑO DE UN MODELO DE DIÁLOGO PARA LA CREACIÓN DE UN AGENTE CONVERSACIONAL

Aplicación en el dominio de la formación digital

ESPECIALIDAD: Colecciones digitales

ALUMNA: Adalía Cuesta Bergstrom

DNI: 02731268H

CONVOCATORIA: Noviembre

TUTORA EN UNIVERSIDAD: Ana María Fernández-Pampillón Cesteros

COTUTORA EN EMPRESA: Isabel Blas Moreno



an NTT DATA Company



**MÁSTER EN: LETRAS DIGITALES: ESTUDIOS AVANZADOS EN
TEXTUALIDADES ELECTRÓNICAS**

**COMPROMISO DEONTOLÓGICO PARA LA ELABORACIÓN, REDACCIÓN
Y POSIBLE PUBLICACIÓN DEL TRABAJO DE FIN DE MÁSTER (TFM)**

**CENTRO: FACULTAD DE FILOLOGÍA, UNIVERSIDAD COMPLUTENSE DE
MADRID**

ESTUDIANTE DE MÁSTER: Adalía Cuesta Bergstrom

TUTOR/ES DEL TFM: Ana María Fernández Pampillón Cesteros

**TÍTULO DEL TFM: Diseño de un modelo de diálogo para la creación de un
agente conversacional. Aplicación en el dominio de la formación digital**

FECHA DE PRIMERA MATRÍCULA: Noviembre de 2018

FECHA DE SEGUNDA MATRÍCULA (en caso de producirse):

1. Objeto

El presente documento constituye un compromiso entre el estudiante matriculado en el Máster en Letras digitales: Estudios Avanzados en Textualidades Electrónicas y su Tutor/es y en el que se fijan las funciones de supervisión del citado trabajo de fin de máster (TFM), los derechos y obligaciones del estudiante y de su/s profesor/es tutor/es del TFM y en donde se especifican el procedimiento de resolución de potenciales conflictos, así como los aspectos relativos a los derechos de propiedad intelectual o industrial que se puedan generar durante el desarrollo de su TFM.

2. Colaboración mutua

El/los tutor/es del TFM y el autor del mismo, en el ámbito de las funciones que a cada uno corresponden, se comprometen a establecer unas condiciones de colaboración que permitan la realización de este trabajo y, finalmente, su defensa de acuerdo con los procedimientos y los plazos que estén establecidos al respecto en la normativa vigente.

3. Normativa



Los firmantes del presente compromiso declaran conocer la normativa vigente reguladora para la realización y defensa de los TFM y aceptan las disposiciones contenidas en la misma.

4. Obligaciones del estudiante de Máster

- Elaborar, consensuado con el/los Tutor/es del TFM un cronograma detallado de trabajo que abarque el tiempo total de realización del mismo hasta su lectura.
- Informar regularmente al Tutor/es del TFM de la evolución de su trabajo, los problemas que se le planteen durante su desarrollo y los resultados obtenidos.
- Seguir las indicaciones que, sobre la realización y seguimiento de las actividades formativas y la labor de investigación, le hagan su tutor/es del TFM.
- Velar por el correcto uso de las instalaciones y materiales que se le faciliten por parte de la Universidad Complutense con el objeto de llevar a cabo su actividad de trabajo, estudio e investigación.

5. Obligaciones del tutor/es del TFM

- Supervisar las actividades formativas que desarrolle el estudiante; así como desempeñar todas las funciones que le sean propias, desde el momento de la aceptación de la tutorización hasta su defensa pública.
- Facilitar al estudiante la orientación y el asesoramiento que necesite.

6. Buenas prácticas

El estudiante y el tutor/es del TFM se comprometen a seguir, en todo momento, prácticas de trabajo seguras, conforme a la legislación actual, incluida la adopción de medidas necesarias en materia de salud, seguridad y prevención de riesgos laborales.

También se comprometen a evitar la copia total o parcial no autorizada de una obra ajena presentándola como propia tanto en el TFM como en las obras o los documentos literarios, científicos o artísticos que se generen como resultado del mismo. Para tal, el estudiante firmará la Declaración de No Plagio del ANEXO I, que será incluido como primera página de su TFM.

7. Procedimiento de resolución de conflictos académicos

En el caso de producirse algún conflicto derivado del incumplimiento de alguno de los extremos a los que se extiende el presente compromiso a lo largo del desarrollo de su TFM, incluyéndose la posibilidad de modificación del nombramiento del tutor/es, la coordinación del máster buscará una solución consensuada que pueda ser aceptada por las partes en conflicto. En ningún caso el estudiante podrá cambiar de



Tutor directamente sin informar a su antiguo Tutor y sin solicitarlo oficialmente a la Coordinación del Máster.

En el caso de que el conflicto persista se gestionará según lo previsto en el SGIC de la memoria verificada.

8. Confidencialidad

El estudiante que desarrolla un TFM dentro de un Grupo de Investigación de la Universidad Complutense, o en una investigación propia del Tutor, que tenga ya una trayectoria demostrada, o utilizando datos de una empresa/organismo o entidad ajenos a la Universidad Complutense de Madrid, se compromete a mantener en secreto todos los datos e informaciones de carácter confidencial que el Tutor/es del TFM o de cualquier otro miembro del equipo investigador en que esté integrado le proporcionen así como a emplear la información obtenida, exclusivamente, en la realización de su TFM.

Asimismo, el estudiante no revelará ni transferirá a terceros, ni siquiera en los casos de cambio en la tutela del TFM, información del trabajo, ni materiales producto de la investigación, propia o del grupo, en que haya participado sin haber obtenido, de forma expresa y por escrito, la autorización correspondiente del anterior Tutor del TFM.

9. Propiedad intelectual e industrial

Cuando la aportación pueda ser considerada original o sustancial el estudiante que ha elaborado el TFM será reconocido como cotitular de los derechos de propiedad intelectual o industrial que le pudieran corresponder de acuerdo con la legislación vigente.

10. Periodo de Vigencia

Este compromiso entrará en vigor en el momento de su firma y finalizará por alguno de los siguientes supuestos:

- Cuando el estudiante haya defendido su TFM.
- Cuando el estudiante sea dado de baja en el Máster en el que fue admitido.
- Cuando el estudiante haya presentado renuncia escrita a continuar su TFM.
- En caso de incumplimiento de alguna de las cláusulas previstas en el presente documento o en la normativa reguladora de los Estudios de Posgrado de la Universidad Complutense.

La superación académica por parte del estudiante no supone la pérdida de los derechos y obligaciones intelectuales que marque la Ley de Propiedad Intelectual para ambas partes, por lo que mantendrá los derechos de propiedad intelectual sobre su trabajo, pero seguirá obligado por el compromiso de confidencialidad respecto a los proyectos e información inédita del tutor.



We

Firmado en Madrid, a 04 de noviembre de 2018

<p>El estudiante de Máster</p> <p style="text-align: center;"><i>Adalía CB</i></p> <p>Fdo.:</p>	<p>El Tutor/es</p> <p>Fdo.:</p>
---	---------------------------------

SR. COORDINADOR DEL MÁSTER EN LETRAS DIGITALES: ESTUDIOS AVANZADOS EN TEXTUALIDADES ELECTRÓNICAS



ANEXO I: DECLARACIÓN DE NO PLAGIO

D./Dña. Adalía Cuesta Bergstrom con NIF 02731268H, estudiante de Máster en la Facultad de Filología de la Universidad Complutense de Madrid en el curso 2017 - 2018, como autor/a del trabajo de fin de máster titulado “Diseño de un modelo de diálogo para la creación de un agente conversacional. Aplicación en el dominio de la formación digital” y presentado para la obtención del título correspondiente, cuyo/s tutor/ es/son:

Ana María Fernández Pampillón Cesteros

DECLARO QUE:

El trabajo de fin de máster que presento está elaborado por mí y es original. No copio, ni utilizo ideas, formulaciones, citas integrales e ilustraciones de cualquier obra, artículo, memoria, o documento (en versión impresa o electrónica), sin mencionar de forma clara y estricta su origen, tanto en el cuerpo del texto como en la bibliografía. Así mismo declaro que los datos son veraces y que no he hecho uso de información no autorizada de cualquier fuente escrita de otra persona o de cualquier otra fuente.

De igual manera, soy plenamente consciente de que el hecho de no respetar estos extremos es objeto de sanciones universitarias y/o de otro orden.

En Madrid, a 04 de noviembre de 2018

A handwritten signature in black ink that reads 'Adalía CB'.

Fdo.:

Esta DECLARACIÓN debe ser insertada en primera página de todos los trabajos fin de máster conducentes a la obtención del Título.

Agradecimientos

Siempre me ha parecido que la parte de más difícil de cualquier trabajo es aquella en la que el autor se expone a sí mismo, porque también es aquella en la que refleja sus propios sentimientos. No es difícil por los contenidos, pues el autor en su mente los tiene bien claros, sino por encontrar la forma más adecuada para plasmarlos; es de vital importancia que las palabras reflejen fielmente lo que se quiere expresar. Por lo tanto, ahora voy a tratar de expresar mis más sinceros agradecimientos a las siguientes personas.

En primer lugar, quiero darle las gracias a la coordinadora del máster, Amelia, por no haber dejado que perdiera las ganas de continuar con este trabajo y por haberme alentado a intentar llegar a esta convocatoria; y a mi tutora del TFM, Ana, por haberme aceptado con un margen de tiempo tan escaso y asumiendo mi tutelado aun con el resto de los trabajos que dirigía.

En segundo lugar, también quiero dar las gracias a Everis por haberme dejado hacer mi trabajo sobre mis prácticas. En especial, quiero agradecerle a mi tutora de las prácticas y del TFM en la empresa, Isabel, por haberme ayudado a crecer como trabajadora y como persona, y por todo el tiempo que me ha dedicado y todas las facilidades que me ha dado; sin ello seguramente no estaría aquí ahora.

En tercer lugar, gracias a mi familia por haberme estado apoyando desde el día que empecé, no solo el máster sino la carrera. Gracias por haberme ayudado a entender que tenía que estudiar aquello que me gustara y no aquello que tuviera más salidas laborales. Gracias a ello he podido comprobar y demostrar que si hacemos lo que nos gusta podemos brillar y encontrar nuestra vocación y el trabajo de nuestros sueños.

En cuarto y último lugar, quiero dar las gracias a mi pareja y a mis amigos, por haber estado disponibles siempre que necesitaba comentar mis dudas o ver la opinión de lo que iba escribiendo. Sobre todo, gracias a todos los que de buena gana os habéis ido leyendo mi trabajo y me habéis ido aconsejando qué cambiar, qué añadir y qué quitar. Sin vosotros este TFM tendría un aspecto completamente diferente.

Índice

Resumen	1
Abstract	1
1. Introducción	3
2. Objetivo del proyecto	6
3. Método de trabajo	7
4. Estado de la cuestión	9
4.1 <i>El Procesamiento del Lenguaje Natural y la representación del conocimiento</i>	<i>10</i>
4.2 <i>Sistemas de diálogo.....</i>	<i>13</i>
4.2.1 <i>Agente conversacional o chatbot</i>	<i>14</i>
4.2.2 <i>Agentes conversacionales basados en marcos</i>	<i>17</i>
4.3 <i>La conversación humana y los agentes conversacionales</i>	<i>18</i>
4.3.1 <i>Teoría de marcos de Minsky.....</i>	<i>19</i>
4.3.2 <i>Teoría de los guiones de Schank y Abelson</i>	<i>20</i>
4.4 <i>Sistemas para representar flujos conversacionales</i>	<i>24</i>
4.5 <i>Agentes conversacionales en el mundo empresarial.....</i>	<i>27</i>
5. Proyecto	30
5.1 <i>Proyecto de Everis</i>	<i>30</i>
5.1.1 <i>Objetivos de Everis.....</i>	<i>30</i>
5.2 <i>Proyecto del TFM</i>	<i>31</i>
5.3 <i>Propuesta de modelo de conversación para el agente</i>	<i>33</i>
5.3.1 <i>Requisitos del modelo de conversación.....</i>	<i>33</i>
5.3.2 <i>Metamodelo del modelo de conversación</i>	<i>35</i>
5.3.3 <i>Propuesta 1: modelo de conversación en forma de flujograma</i>	<i>36</i>
5.3.4 <i>Propuesta 2: modelo de conversación en forma de guion “literario”</i>	<i>37</i>
6. Validación y pruebas	38
6.1 <i>Encuesta.....</i>	<i>38</i>
6.2 <i>Resultados obtenidos.....</i>	<i>38</i>
6.3 <i>Discusión de resultados</i>	<i>43</i>

7. Resumen, conclusiones y trabajo futuro	44
8. Referencias	46
9. Anexos.....	49

Índice de figuras

Figura 1. Aplicaciones de mensajería instantánea más utilizadas en octubre de 2018	5
Figura 2. Diagrama de Gantt	7
Figura 3. Acercamientos a la comprensión del lenguaje natural	11
Figura 4. Etapas de la comprensión del lenguaje natural	11
Figura 5. Ejemplo de conversación de Cobot.....	17
Figura 6. Ejemplo del modelo seq2seq2.....	17
Figura 7. Ejemplo de marco en GUS.....	18
Figura 8. Estructura general de un marco.....	20
Figura 9. Marco específico sobre el animal león.....	20
Figura 10. Guion del restaurante parte I.....	23
Figura 11. Guion del restaurante parte II.....	24
Figura 12. Flujograma representando una conversación.....	26
Figura 13. Flujograma representando una conversación.....	27
Figura 14. Guion representando una conversación	27
Figura 15. Ejemplo de guion para el caso de uso 1	36
Figura 16. Gráfico de los resultados de la pregunta 1	39
Figura 17. Gráfico de los resultados de la pregunta 2	39
Figura 18. Gráfico de los resultados de la pregunta 3	39
Figura 19. Gráfico de los resultados de la pregunta 4	40
Figura 20. Gráfico de los resultados de la pregunta 5	40
Figura 21. Gráfico de los resultados de la pregunta 7	41

Resumen

Los agentes conversacionales o chatbots han adquirido un papel importante últimamente en el mundo tecnológico, pese a que los sistemas de diálogo llevan existiendo desde mediados de siglo pasado. Esto ocurre porque las empresas han comenzado a necesitar mostrar una mayor y mejor comunicación con sus clientes o potenciales clientes y, en este sentido, los agentes conversacionales proporcionan esta comunicación (o sensación de comunicación) a un coste razonable.

Para conseguir que los agentes conversacionales cumplan con las expectativas que se depositan en ellos es crucial realizar con la máxima precisión y completitud posible la tarea de la modelización del flujo conversacional que es donde se define el patrón de las conversaciones van a mantener el usuario con el agente. Precisamente en esta tarea cobra especial importancia la figura del lingüista, pues es el único perfil profesional preparado para poder diseñar estos flujos de manera que se obtengan los resultados más satisfactorios desde el punto de vista de la experiencia de usuario: cuanto mejor esté definido, más natural y humana parecerá la conversación.

En este trabajo, se pone el foco en la búsqueda de un modelo que represente de forma precisa y formal las conversaciones que mantendrá un agente conversacional, llamado Lía, con los usuarios que necesitan buscar formación en línea. De esta búsqueda resulta, no sólo el modelo de conversación sino también el desarrollo de un sistema o metamodelo que mejora la usabilidad de los sistemas tradicionalmente utilizados. Este metamodelo ha sido evaluado empíricamente por los especialistas en informática que lo han usado para el agente Lía. En este sentido, los resultados obtenidos pueden servir para facilitar el diseño de nuevos modelos de conversación para este agente u otros con aplicación a nuevos dominios de conocimiento.

Palabras clave: agentes conversacionales, sistemas de diálogo, chatbot, Inteligencia Artificial, Procesamiento del Lenguaje Natural, representación del conocimiento, flujos conversacionales

Abstract

In the past recent years, conversational agents or chatbots have acquired a major role in the technological world, despite they have been a reality since the last half of the 20th century. This is because companies have started feel the need to show a better communication with their clients or potential clients and, in that sense, conversational

agents provide this communication – or at least they give this sensation – at a reasonable cost.

To get these agents to meet the expectations people put in them it is crucial to perform with the maximum precision and completeness the task of modeling the conversational flow, because it is where the pattern of the conversations between the conversational agent and its users is defined. As a matter of fact, in this task there is one profile that acquires special significance: the linguist, since it is the only professional prepared to design these flows to achieve the most satisfying results from a user experience point of view: the better it is defined, the more natural it will seem.

In this paper, focus is put in the search of a model that precisely and formally represents the conversations that a conversational agent, called Lía, will have with the users that need to find online training. The results of this search are, not only the conversational model, but also the development of a system or metamodel which improves the usability of the traditionally used systems. This metamodel has been empirically evaluated by the software specialists that have used it with agent Lía. In this sense, the results that have been obtained can serve to improve the design of new models of conversation for this agent or for others, and in new areas of knowledge.

Since the rediscovery of the chatbots in the last few years, there have been many consultancy firms that have started to develop these kinds of software programs to be able to cover the new necessities that have arisen in the other companies. In our case, we are going to focus in finding a model that represents the conversation a conversational agent has with its users in the online course search. In this manner, we will obtain an efficient conversational flow model that can be easily and visually understood by the software developers.

Key words: conversational assistant, dialog systems, chatbot, artificial intelligence, natural language processing, knowledge representation, conversational flow

1. Introducción

El ser humano es, como bien indica Aristóteles, un *zoon politikon*, es decir, un animal social o político, que se diferencia del resto de seres por su capacidad de relacionarse socialmente con otros y de crear sociedades. Esta naturaleza del hombre y de la mujer ocasiona que la comunicación cobre especial importancia, pues es vital para permitir esa unión entre personas y permitir que se adscriban a diferentes grupos que acaban configurando las diferentes sociedades que se han dado a lo largo de la historia. El elemento clave que permite que nos comuniquemos entre nosotros es el lenguaje y, por ello, es uno de los rasgos más característicos de nuestra especie: lo aprendemos desde que nacemos y lo utilizamos a lo largo de toda nuestra vida para poder conversar con otras personas y, así, ir creando nuestras tribus sociales dentro del marco global de la sociedad en la que vivimos.

Por tanto, queda patente que el lenguaje es el que permite que cumplamos uno de los objetivos más importantes para la humanidad: sentirse reconocido como parte de un grupo, establecer y mantener lazos sociales con sus integrantes. En este sentido, la conversación humana es una de las posibilidades clave del lenguaje puesto que, a través de ella, una persona consigue desenvolverse como un miembro funcional y activo dentro de una sociedad.

En torno a la conversación, cobra cierta importancia una pregunta clave que nos hemos formulado a lo largo de toda la historia: ¿cómo organizamos el conocimiento de forma que sabemos qué decir y qué hacer en cada momento? De este tema, en el área de la Lingüística, se encarga principalmente la Pragmática.

En los últimos años, la idea de conseguir que las máquinas se comuniquen con los seres humanos ha pasado de la ciencia ficción a la realidad y, por tanto, gran parte de las cuestiones que estudiaban los lingüistas han pasado a ser cuestiones que necesitan resolver los informáticos. Entre ellas destaca la pregunta ¿es posible que un ordenador mantenga una conversación con una persona o con otro ordenador de la misma forma que la mantendría una persona?, y si es así, ¿cómo?

Desde mediados del siglo pasado, surgió un nuevo campo de estudio conocido como Inteligencia Artificial (IA) y, con ella, del Procesamiento del Lenguaje Natural (PLN) y los sistemas de diálogo. A raíz de esto, la comunidad científica comenzó a darle mayor importancia a temas nunca considerados, como la forma con la que se podría

conseguir que un robot organizara la información de forma similar a los seres humanos para que pudieran comunicarse con nosotros.

Puede parecer que los agentes conversacionales son una invención del siglo XXI, la verdad es que llevan existiendo desde la segunda mitad del siglo XX: «El primer chatbot fue introducido incluso antes de que se creara el primer ordenador personal. Se llamó Eliza y fue desarrollado en el Laboratorio de Inteligencia Artificial del MIT por Joseph Weizenbaum en 1966.»¹ (Khan y Das, 2018, p. 2). No obstante, aunque se siguió avanzando en este campo con el desarrollo de otros agentes conversacionales, no es hasta la llegada de la era de los teléfonos inteligentes (o *smartphones*), en 2007 con el lanzamiento del primer iPhone (Arthur, 2012), que comienzan a tomar fuerza en el campo comercial, tal y como queda reflejado en el anexo I. De hecho, fue Siri –el agente nativo de Apple– el primero de los agentes desarrollados por empresas privadas para vender a sus clientes en 2010. Unido a esto, llegó el auge de las plataformas de mensajería instantánea como Facebook Messenger, WhatsApp o Telegram, que facilitaban, entre otras cosas, la conexión directa entre clientes y empresas (Visual Capitalist, 2016).

Este Trabajo de Fin de Máster (TFM) se enmarca en este contexto de creación de agentes conversacionales también llamados, simplemente, agentes. El TFM se plantea como parte de un proyecto de creación de un agente que surge tras el encargo de una empresa tecnológica a la empresa Everis². Esto es debido a que el sector empresarial actual está buscando nuevas formas de conectar con sus clientes y, como la sociedad actual vive en un mundo digital y se comunica principalmente a partir de plataformas de mensajería instantáneas –ver figura 1, obtenida de Statista (2018)–, el agente conversacional es la mejor vía para poderse acercar a su público.

¹ Traducción propia del original: «Even though *chatbot* seems to be a recent buzzword, they've been in existence since people developed a way to interact with computers. The first-ever chatbot was introduced even before the first personal computer was developed. It was named Eliza and was developed at the MIT Artificial Intelligence Laboratory by Joseph Weizenbaum in 1966».

² Everis an NTTT Data Company es una empresa multinacional que ofrece servicios de consultoría de aplicaciones tecnológicas y outsourcing en todos los sectores del ámbito económico.

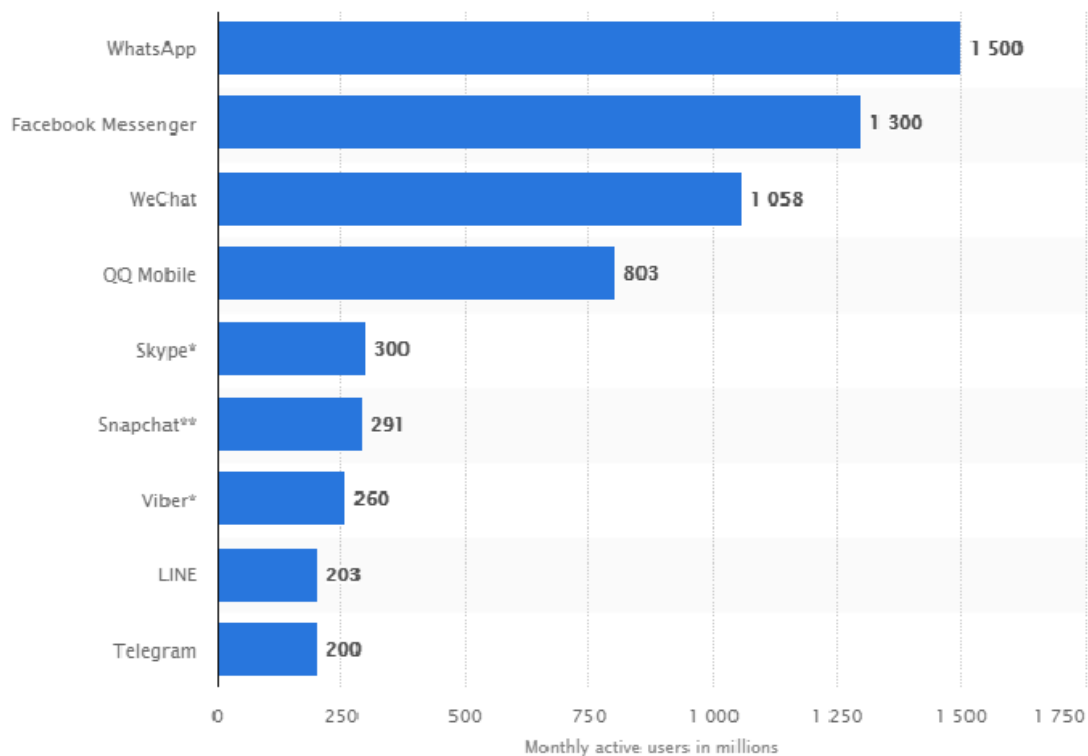


Figura 1. Aplicaciones de mensajería instantánea más utilizadas en octubre de 2018

El objetivo del encargo era buscar la mejor manera de crear un orientador que ayude y asesore a personas en situación de desempleo o con la intención de cambiar su orientación laboral. Para ello, el agente primero recomienda al usuario una serie de carreras profesionales en las que puede encajar, y luego le muestra diferentes cursos de plataformas online que le permitan formarse en las destrezas técnicas que requiere la carrera o carreras en las que se interese.

Dado que la autora de este documento trabaja en esta consultora como estudiante en prácticas y se encarga de conceptualizar y crear los flujos conversacionales de diferentes agentes, la cuestión inicial en torno a la que se centra mi proyecto es la siguiente: ¿cuál es el modelo de conversación adecuado para un agente conversacional en el dominio de la búsqueda de cursos online?

Encontrar un buen modelo conversacional significa conseguir representar la conversación adecuadamente de manera que, mediante esa representación, se entienda correctamente la funcionalidad que debe desarrollarse para que el intercambio comunicativo entre el agente y el usuario resulte satisfactorio. El modelo debe ser capaz de abarcar las situaciones de diálogo entre el agente y cualquier usuario e interactuar con este de modo que sienta que está interactuando con un ente, persona o máquina, inteligente que le ayudará en lo que necesite.

2. Objetivo del proyecto

El objetivo principal de este proyecto es diseñar un modelo de conversación dentro del dominio de conocimiento de la búsqueda de cursos online para el agente conversacional creado la empresa Everis (en adelante, el agente conversacional Everis o agente Everis) para un cliente particular. El modelo debe representarse, además, en un formato fácil de entender por parte de los desarrolladores informáticos del agente conversacional.

Este objetivo se puede descomponer en las siguientes metas específicas:

- Análisis del estado de la cuestión para conocer los modelos conversacionales ya existentes. A raíz de esto, se contextualizará el proyecto para justificar el uso de unos modelos de representación de flujos conversacionales y sobre otros modelos y cuáles son los más usables.
- Selección del formato del modelo o los modelos (metamodelos) que mejor representen los flujos conversacionales del agente Everis.
- Desarrollo del modelo o modelos seleccionados para el agente de Everis, en concreto, a un caso de uso dentro del ámbito en el que se enfoca el proyecto.
- Evaluación de la usabilidad del modelo o modelos desarrollados.

3. Método de trabajo

Para poder cumplir el objetivo de proyecto y, por tanto, dar respuesta a la cuestión inicial, se han definido cinco fases principales, cada una con sus correspondientes tareas necesarias, como se muestra en la figura 2.

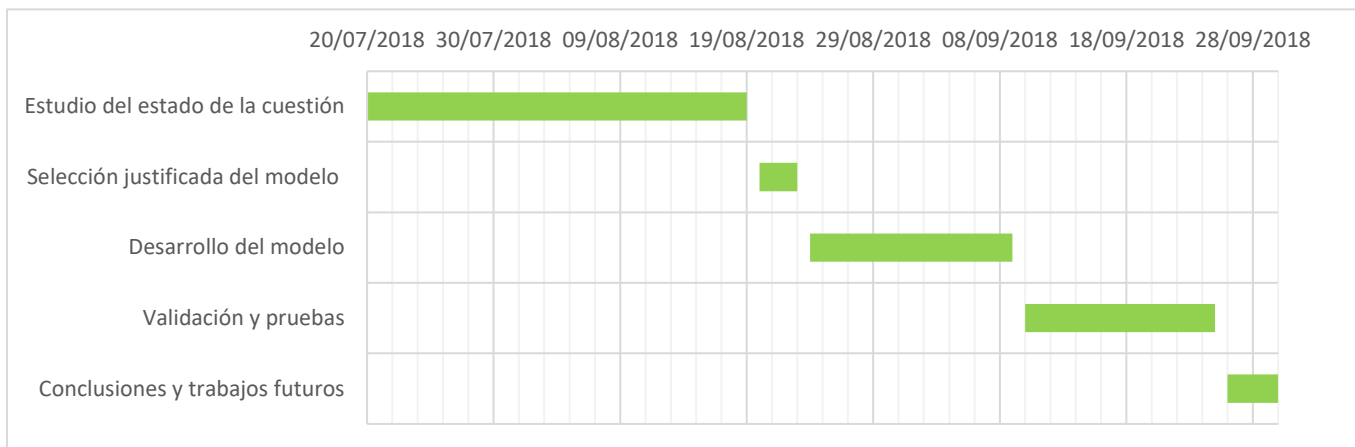


Figura 2. Diagrama de Gantt

Primera fase: estudio del estado de la cuestión

En esta etapa nos centramos en adquirir toda la información necesaria para poder tener una visión global de la generación del modelo de diálogo en agentes conversacionales. En concreto, por tanto, debemos investigar sobre la Inteligencia Artificial, la representación del conocimiento, el Procesamiento del Lenguaje Natural, los sistemas de diálogo, la importancia de la pragmática y el contexto en la conversación humana y las herramientas que existen actualmente para diseñar agentes conversacionales.

Una vez recopilada la información relevante para el TFM, se debe realizar un análisis detallado y en profundidad de esta bibliografía recopilada para ver qué sirve y en qué medida.

El último paso es la síntesis del conocimiento adquirido y la redacción del estado de la cuestión, donde quedará reflejado todo este proceso.

Segunda fase: selección justificada del modelo o modelos

Tras la fase de documentación, se procederá a realizar una selección justificada del modelo o de los modelos de diálogo que se crean más adecuados para la representación de la conversación entre un humano y una máquina. Con esto se persigue establecer cuáles son las mejores maneras de representar un modelo de diálogo eficaz para un agente conversacional dentro del dominio tecnológico en el que nos encontramos.

Tercera fase: desarrollo del proyecto

Después de la selección del modelo o modelos de diálogo que se consideran más adecuados estos se aplican al diseño y construcción del modelo de diálogo específico del agente Everis. El modelo de diálogo creado se integrará en el agente conversacional por parte de los desarrolladores informáticos del mismo.

Cuarta fase: validación y pruebas

Se evalúa la usabilidad del modelo de diálogo Everis construido en la fase anterior e integrado en el agente conversacional. Se debe verificar que el agente puede mantener conversaciones con los usuarios de la forma esperada.

Quinta fase: conclusiones y trabajos futuros

Con los resultados de la validación se elaborarán las conclusiones sobre la usabilidad del modelo de diálogo desarrollado en este TFM y sobre el propio TFM. Además, se propondrán una serie de trabajos futuros que pueden surgir a partir de este proyecto.

4. Estado de la cuestión

La Inteligencia Artificial –entendida como «la rama de las ciencias computacionales que se ocupa de diseñar sistemas informáticos inteligentes, es decir, sistemas que muestren características que asociamos con la inteligencia propia del comportamiento humano: entendimiento del lenguaje, aprendizaje, razonamiento, resolución de problemas, etc.» (Barr y Feigenbaum, 1981, p. 8)³– surgió con la llegada de los primeros computadores en el siglo pasado y, gracias a varios científicos de mediados de siglo como Alan Turing o John McCarthy, se estableció como la rama del saber que es en la actualidad.

Definir qué es inteligencia no es una tarea sencilla. En la década de 1950 Alan Turing propuso que un sistema inteligente debería tener, como mínimo, las capacidades de (García Serrano, 2016): Procesamiento del Lenguaje Natural (PLN), Razonamiento y Aprendizaje y Representación del Conocimiento. Estos requerimientos, por su magnitud, acabaron convirtiéndose en campos de estudio diferentes de la Inteligencia Artificial.

Para este trabajo, el campo más importante es el PLN. Una aplicación o sistema PLN es una aplicación software capaz de entender o analizar y/o producir o sintetizar Lenguaje Natural (LN) (Allen, 1995), que frente al artificial es el lenguaje que utilizan las personas para comunicarse. Una de las aplicaciones del PLN son las interfaces en Lenguaje Natural denominadas *sistemas de diálogo* o *agentes conversacionales*. Un agente conversacional es «un sistema automático capaz de emular a un ser humano en un diálogo con otra persona, con el objetivo de que el sistema cumpla con una cierta tarea (normalmente suministrar una cierta información o llevar a cabo una determinada tarea» (Griol, Zoraida, López-Cózar y Gutiérrez, 2009, p. 2).

En este trabajo, nos proponemos encontrar un modelo para representar el conocimiento que debe tener un agente conversacional sobre un dominio específico para poder procesar LN, es decir, para poder mantener, de forma eficaz, conversaciones con seres humanos.

Por ello se presenta una síntesis, en las tres secciones siguientes, de los modelos de representación del conocimiento lingüístico desde el punto de vista del PLN, de los agentes conversacionales, y, de la teoría del diálogo desde el punto de vista Lingüístico

³ Traducción propia del original: «...the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behavior: understanding language, learning, reasoning, solving problems, and so on.».

de la Pragmática. Esta síntesis servirá de base para justificar la selección del modelo de representación del diálogo que se ha creado en este TFM.

4.1 El Procesamiento del Lenguaje Natural y la representación del conocimiento

El Procesamiento del Lenguaje Natural es un área de estudio dentro de la Inteligencia Artificial que se encarga de estudiar el uso de métodos que permitan a las computadoras aprender, entender y producir contenido en lenguaje natural (Hirschberg y Manning, 2015). Dado su amplio alcance, los sistemas de Procesamiento del Lenguaje Natural (en adelante SPLN) pueden tener varias aplicaciones: asistir en la comunicación entre seres humanos con la traducción automática, ayudar en la conversación humano-máquina con agentes conversacionales, o beneficiar tanto a las personas como a las computadoras mediante el análisis y el aprendizaje de la gran cantidad de contenido de lenguaje natural que hay disponible en la web.

En general un SPLN toma una forma del lenguaje –ya sea un diálogo hablado o un texto escrito– y extrae su significado. Un agente conversacional es un SPLN, ya que es necesario que entienda qué quiere el usuario. Así, por ejemplo, si una persona pide a su agente que le ponga un recordatorio a las siete de la tarde para que llame a su madre, el agente debe entender cada una de las piezas del mensaje para realizar bien su trabajo: tipo de acción, agente, objeto, tiempo y beneficiario de la acción.

El mensaje en lenguaje natural puede proceder de muchas fuentes diferentes: conversaciones con agentes humanos, tuits, correos, wasaps, facturas... incluso cartas escritas a mano; el objetivo es encontrar el significado del mensaje. Existen tres tipos diferentes de acercamientos a la comprensión del lenguaje natural, como se observa en la figura 3⁴ (Dahl, 2013):

- **ANÁLISIS TEXTUAL:** la unidad básica de análisis es el texto en sí mismo. Se aplican, fundamentalmente, modelos del lenguaje estadísticos como la proximidad de unas palabras con otras o la frecuencia relativa de ciertas palabras en diferentes textos.
- **ANÁLISIS SEMÁNTICO-SINTÁCTICO:** en este tipo de análisis se intenta buscar algún tipo de estructura dentro del texto para obtener información específica. Por tanto, no se estudia el texto en sí mismo, sino las oraciones que lo componen para encontrar estructuras que puedan ser utilizadas por una aplicación.

⁴ Figura traducida de la original disponible en Dahl (2013).

- **ANÁLISIS MULTINIVEL:** consiste en un tipo de análisis muy detallado de las aportaciones lingüísticas que reciben que se basa en diferentes subcampos de la Lingüística como la Lexicología, la Sintaxis, la Semántica y la Pragmática (Dahl, 2013). Cada campo tiene una importancia relativa según el lenguaje. Por ejemplo, en las lenguas aislantes (lenguas en las que las palabras son invariables), como el chino (Sapir, 1921/1980), cobra mayor importancia la pragmática pues no se hacen tantas especificaciones sintácticas como en el resto de lenguas y el significado se saca por medio del contexto en el medio escrito.

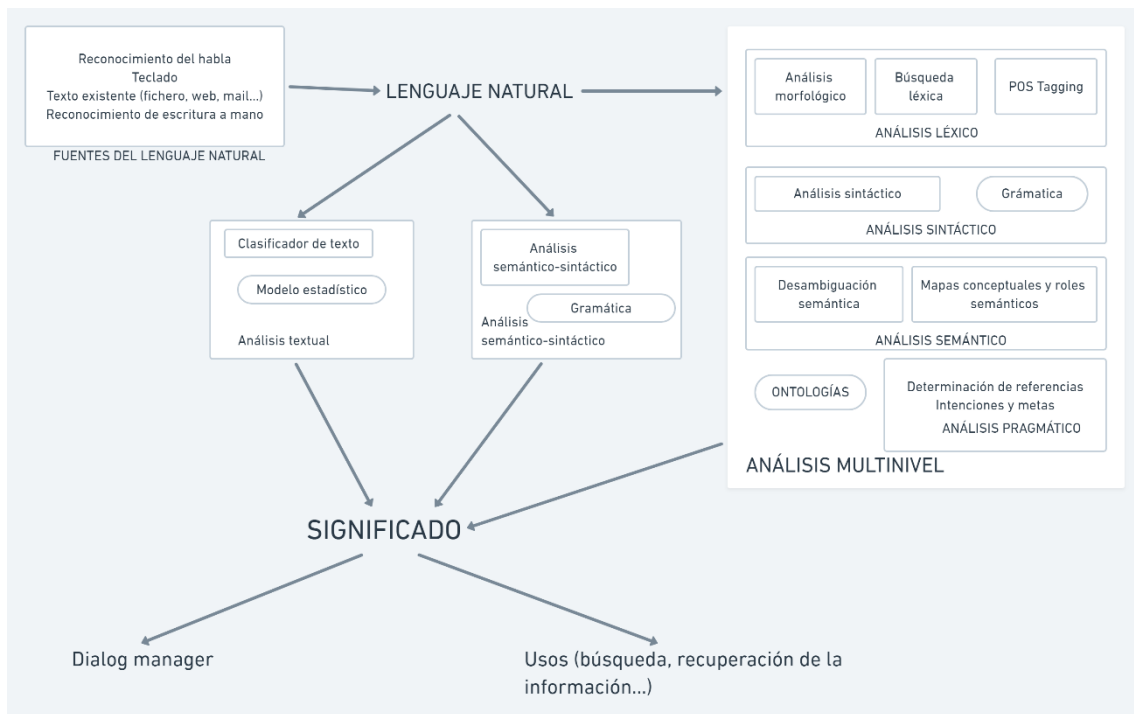


Figura 3. Acercamientos a la comprensión del lenguaje natural

De acuerdo con Cortez Vásquez, Vega Huerta y Paquinona Quispe (2009), un sistema de Procesamiento del Lenguaje Natural sigue, en general, tres etapas (figura 4):

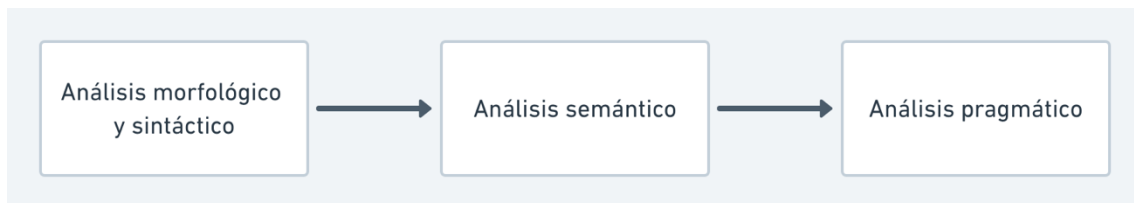


Figura 4. Etapas de la comprensión del lenguaje natural

1. **ANÁLISIS MORFOLÓGICO Y SINTÁCTICO.** El ordenador analiza las oraciones recibidas del usuario a nivel morfológico y sintáctico para obtener una representación morfosintáctica de la oración. Aquí tienen gran importancia el analizador léxico, que identifica los componentes léxicos etiquetándolos con su categoría gramatical y,

posiblemente, lematizándolos, y, el analizador sintáctico, que comprueba si se cumple el orden gramatical entre los elementos identificados.

2. ANÁLISIS SEMÁNTICO. A partir de la representación morfosintáctica se calcula una representación semántica expresada, en la mayoría de los sistemas como una expresión lógica.
3. ANÁLISIS PRAGMÁTICO. Se analizan la secuencia de representaciones semánticas de las frases para obtener una representación final del conocimiento de la entrada del usuario en un formalismo adecuado al procesamiento que se necesite realizar posteriormente con esta representación (acceso a una base de datos, inferencias, etc.)

La obtención de una representación del conocimiento a nivel pragmático es la etapa más compleja del PLN y la que, actualmente, ofrece peores resultados. La razón es la cantidad de conocimiento que se necesita, además del conocimiento lingüístico para una interpretación final, adecuada con el uso que los hablantes hacen del lenguaje en la comunicación. Es decir, la interpretación final de una producción en lenguaje natural depende de las condiciones que determinan el empleo de enunciados concretos emitidos por hablantes concretos en situaciones comunicativas concretas, y su interpretación por parte de los destinatarios (Escandell Vidal, 2013). Entre estas condiciones se pueden destacar dos:

- El **contexto**, que como bien explica Portolés (2004, p. 99) «es siempre mental y lo forman un conjunto de suposiciones que permiten la comprensión de un enunciado, estas suposiciones o se hallan ya en nuestra memoria, o se crean en nuestra mente en el momento de la comunicación». El contexto es vital para relacionar el conocimiento antiguo (el que tenemos) con el nuevo (el que recibimos) y de esta forma poder organizarlo de forma coherente con el conocimiento antiguo.
- El **lenguaje humano** se basa en la capacidad de comunicarse mediante signos lingüísticos y de entender lo que se quiere transmitir aplicando el contexto en el que tiene lugar la conversación. Se caracteriza porque la comunicación humana es: ostensiva, pues la lengua es un estímulo en sí misma y si alguien se comunica con otra persona se debe a que quiere transmitirle algo y el destinatario, automáticamente, lo interpreta; e inferencial, porque dentro de un contexto, el ser humano es capaz de representar lo que quiere decir el interlocutor.

Por tanto, las palabras son el estímulo y, cuando se relacionan con el contexto, se produce el acto de inferencia para averiguar qué significa lo dicho en ese contexto. Así pues, toda comunicación humana es contextual. La capacidad de integrar el contexto o

inferir se ha descrito con diferentes teorías pragmáticas, entre las que destacan, desde el punto de vista de la Inteligencia Artificial, la teoría de los marcos de Minsky o la teoría de los guiones de Abelson y Schank, de las que se hablará más adelante.

Como se verá más adelante, uno de los principales objetivos de los agentes conversacionales es poder mantener una conversación con un humano y realizar con éxito las tareas que se le vayan proponiendo. Para ello, es importante que el agente procese correctamente el lenguaje natural y que, además, tenga un buen sistema de representación del conocimiento para que pueda extraer inferencias adecuadamente y así, obtenga conclusiones en base a las que actuar.

4.2 Sistemas de diálogo

Los agentes conversacionales, también conocidos como sistemas de diálogo, son programas que se comunican con los usuarios mediante el lenguaje natural y pueden dividirse en dos categorías (Jurafsky y Martin, 2017): los agentes conversacionales o *chatbots* y los agentes conversacionales orientados a tareas.

En los últimos años han experimentado un incremento en todo el mundo por las numerosas aplicaciones que, desde el punto de vista comercial, tienen estos programas como:

- LA ATENCIÓN AL CLIENTE: existen ya numerosos agentes orientados a la atención al cliente, sobre todo enfocados a la resolución de preguntas frecuentes como Irene de Renfe⁵, disponible en su página web o el agente del parque Warner⁶, que puede utilizarse desde Facebook Messenger.
- LA SALUD: en concreto en el campo de la psicología, por ejemplo, en el 2017 apareció *Woebot*⁷ en la plataforma de Facebook Messenger, de momento solo disponible en inglés. Este agente ha sido diseñado empleando terapias cognitivo-conductuales para que los usuarios sean capaces de controlar su estado anímico y, en última instancia, para decirle si está o no deprimido (Blasco, 2017).
- EL ENTRETENIMIENTO: *Ometrics Joke Bot*, un agente que cuenta chistes y que también acepta que los usuarios envíen sus propios chistes.

⁵ <http://consulta.renfe.com/base/main.jsessionid=2B1B01080EBF8D2EB77B9346739189D4>

⁶ <https://www.parquewarner.com/blog/descubre-nuestro-nuevo-chatbot-y-planifica-tu-visita-a-parque-warner>

⁷ <https://woebot.io/>

- **PROMOCIONAR UN PRODUCTO:** cada vez son más las empresas que utilizan a un agente para vender sus productos, en concreto son las propias compañías desarrolladoras de agentes las que los incluyen en sus páginas web para vender al cliente el producto de forma muy visual e interactiva. Un ejemplo de ello puede ser *Specter*⁸ el agente de Maruti Tech Labs.

Actualmente, la mayoría de las empresas dedicadas a la creación de sistemas de diálogo llaman *chatbot* a ambos tipos de agentes conversacionales, pero la mayoría de los que ya están en el mercado, excluyendo a agentes como Siri o Google Assistant, son, en realidad, agentes conversacionales.

4.2.1 Agente conversacional o chatbot

Los agentes conversacionales o *chatbots* son «sistemas que pueden mantener conversaciones extendidas con el objetivo de imitar la desestructuración característica de la interacción entre humanos»⁹ (Jurafsky y Martin, 2017). Por este motivo, suelen incluir un valor de entretenimiento. Un buen ejemplo podría ser *Billybot*¹⁰, el agente de Billy Seguros, que, aunque está diseñado para que los usuarios encuentren el seguro más barato según sus condiciones, incluye un lenguaje muy cercano e informal, además de gifs y emoticonos que le dotan mucha personalidad y, además, entretienen al usuario. Ahora bien, también se han utilizado por motivos prácticos, como probar teorías de terapia psicológica como ELIZA, el primer sistema de diálogo que fue creado por Joseph Weizenbaum en 1966 (Weizenbaum, 1966).

Dentro de esta categoría, podemos encontrar dos clases de agentes conversacionales (Jurafsky y Martin, 2017): los que se basan en reglas y los basados en corpus.

Agentes basados en reglas

Este tipo de agentes sigue un guion conversacional determinado previamente y suelen limitarse al entendimiento de oraciones. Para ello reconocen una serie de construcciones sintácticas específicas, que aprenden gracias a un entrenamiento en el que una o varias personas van nutriéndolo de frases de ejemplo a partir de las cuales puede realizar este reconocimiento.

⁸ <https://www.marutitech.com/>

⁹ Traducción propia del original: «...are systems that can carry on extended conversations with the goal of mimicking the unstructured conversational or ‘chats’ characteristic of human-human interaction. »

¹⁰ <https://billyseguros.com/>

El primer agente conversacional de este tipo que apareció fue ELIZA en 1966. Fue creada por Joseph Weizenbaum y se diseñó para imitar a un psicólogo *rogeriano*. Este tipo de terapia emplea un enfoque no autoritario, que permite que el paciente lleve el mando de la conversación para que llegue por sí mismo a una solución a través de la conversación; el terapeuta se limita a escuchar sin emitir ningún tipo de juicio, aceptando lo que dice el paciente pero sin dirigir la conversación en otras direcciones (Psychology Today, 2018).

ELIZA se basa en el uso de un conjunto de patrones de conversación predeterminados. El sistema busca hacer corresponder la entrada o mensaje del paciente con un patrón. Este patrón de entrada está asociado a uno o más patrones de respuesta del sistema que se utilizan según una estrategia también predeterminada. Para ilustrar mejor esto, podemos utilizar un extracto de una de las conversaciones más famosas de ELIZA, donde las frases en mayúscula son las respuestas del agente (Weizenbaum, 1966):

Men are all alike.

IN WHAT WAY

They're always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE

Well my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME HERE

He says I'm depressed much of the time.

I AM SORRY TO HEAR YOU ARE DEPRESSED

En general, la estructura de la conversación es la siguiente: el texto de entrada se analiza y se busca la presencia de alguna palabra clave, en el ejemplo, la palabra clave es *always*. Si esta se encuentra, se genera una oración en base a una regla asociada a ella; si no, se genera una respuesta sin contenido específico o se recupera una de una parte anterior de la conversación (Weizenbaum, 1966).

Los agentes basados en reglas son muy eficaces para realizar tareas simples, pero, realmente, no entienden al usuario.

Agentes basados en corpus

Este tipo de agentes utilizan minería de textos en diferentes corpus de conversaciones entre humanos (Jurafsky y Martin, 2017) para deducir las reglas, es decir, para extraer patrones estadísticos, en vez de utilizar unas preprogramadas como el tipo de agente anterior. Estos corpus pueden estar compuestos por conversaciones en plataformas de mensajería instantánea (WhatsApp, Messenger, Telegram...), por tuits, por guiones de

películas..., ya que se acercan mucho a la conversación natural y están disponibles en grandes cantidades.

Al igual que los agentes basados en reglas, la mayoría de este tipo de agentes no dan mucha importancia al contexto conversacional, más bien tienden a centrarse en crear una respuesta adecuada para la oración del usuario. Por ello, también son conocidos como sistemas de generación de respuestas.

Existen dos tipos diferentes de este agente conversacional (Jurafsky y Martin, 2017): los que se basan en la recuperación de la información y los que se basan en un aprendizaje automático revisado fundamentado en la transducción de secuencias.

Los *agentes basados en la extracción de información* obtienen los datos necesarios de conversaciones entre seres humanos o de las respuestas que da una persona en una conversación con una computadora. De esta manera, a partir de una oración del usuario seleccionan una respuesta de su corpus que se adapta al tema de la conversación.

Un ejemplo de agente de este tipo sería *Cobot* (Isbell et al., 2006), un agente conversacional que «vivía» dentro de una red social llamada *LambdaMOO*. Como era un agente cuyo objetivo principal era relacionarse con otros usuarios, sus creadores hicieron que extrajera información de ciertos corpus que cumplieran con los siguientes requisitos (Isbell et al., 2006):

- Las oraciones podían servir como frases que diría Cobot, por lo menos en ciertos casos
- Los documentos debían ser lo suficientemente grandes como para que la probabilidad de que se repitan oraciones sea mínima
- Los documentos debían tratar temas específicos y reconocibles
- Las frases generadas por estos documentos deben ser entretenidas y, en ocasiones, controvertidas (para dar pie a que la conversación continúe).

Así pues, los corpus seleccionados fueron *La sociedad industrial y su futuro* de Theodore Kaczynski, un terrorista estadounidense conocido como Unabomber, varios artículos sobre abducciones alienígenas y los guiones de las películas de *El gran Lebowski* y *El planeta de los simios*. En la figura que aparece a continuación se puede observar una conversación entre *Cobot* y un usuario en la que el agente extrae la información del libro escrito por Theodore Kaczynsky:

cobot [to DragonBoi]: The line of conflict should be drawn between the mass of the people and the power-holding elite of industrial society.
 DragonBoi eyes cobot warily.
 cobot [to DragonBoi]: You have something cute in your eye.
 DragonBoi [to HFh]: you're bot is turning into a communist.
 DragonBoi [to cobot]: so what are you gonna do about it?
 DragonBoi [to HFh]: did you see what it told me? you are turning it into a communist!
 Wikked [to cobot]: are you a communist?
 DragonBoi [to cobot]: are you now, or ever been a member of the communist party?
 cobot mumbles something in binary.
 DragonBoi [to cobot]: so you're taking the fifth are you? well this is a special commision, that won't work here!

Figura 5. Ejemplo de conversación de Cobot

Los agentes basados en la codificación y decodificación (en inglés, *sequence to sequence* o *seq2seq*) se basan en la idea de que la generación de respuestas se puede obtener convirtiendo una intervención del usuario en una del agente. (Jurafsky y Martin, 2017). En la siguiente figura, tomada de (Shang, Lu y Li, 2015, p. 1578) se muestra un ejemplo del funcionamiento de este tipo de agentes, en donde se codifica una intervención del usuario para decodificarla en la respuesta del agente.

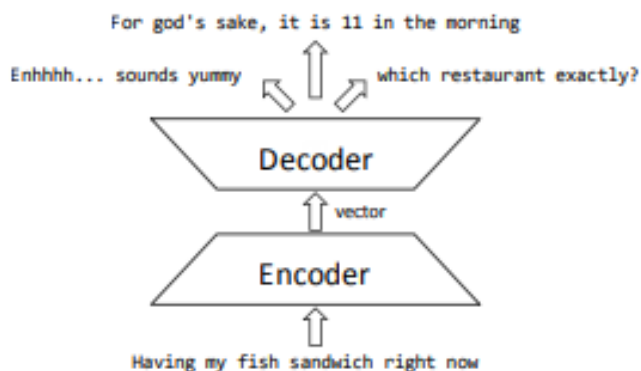


Figura 6. Ejemplo del modelo seq2seq2

4.2.2 Agentes conversacionales basados en marcos

Este tipo de agentes están basados en la teoría de marcos que se presentará en la sección siguiente. Los sistemas de marcos son estructuras de conocimiento que representan los tipos de intenciones que el sistema puede extraer de las oraciones del usuario. Estos sistemas están compuestos de uno o más marcos, entendidos en este contexto como «una estructura de datos que potencialmente contiene un nombre, una referencia a un marco prototipo y un conjunto de ranuras» (Bobrow et al., 1977, p. 163). Uno de los ejemplos más citados para poder explicar este modelo es el del sistema GUS (Genial Understander System) por Bobrow et al. (1977), que fue diseñado

específicamente con el objetivo de planear un viaje dentro del dominio de los viajes. El siguiente ejemplo, tomado de (Bobrow et al., 1977) ilustra un tipo de marco en GUS, en concreto, el relativo a la «fecha» en la que tendrá lugar el viaje. Como se observa, las ranuras de este marco son: mes, día, año, día de la semana y sumario; y cada una de ellas toma unos valores determinados (en el caso de mes el valor es nombre).

```
[DATE
  MONTH      NAME
  DAY        (BOUNDED-INTEGER 1 31)
  YEAR       INTEGER (TOFILL ASSUME 1975)
  WEEKDAY    (MEMBER (SUNDAY MONDAY TUESDAY
                    WEDNESDAY THURSDAY FRIDAY SATURDAY))
             (WHENFILLED FINDDATEFROMDAY)
             (TOFILL GETWEEKDAY))
  SUMMARY (OR (LIST MONTH DAY) WEEKDAY))]
```

Figura 7. Ejemplo de marco en GUS

Este sistema ha sido tan importante que continúa teniendo vigencia en la actualidad pues por ejemplo el agente conversacional de Apple, *Siri*, está basado en el sistema GUS (Jurafsky y Martin, 2017).

4.3 La conversación humana y los agentes conversacionales

La ciencia que se encarga de estudiar la comunicación humana es, esencialmente, la Pragmática, definida por Escandell Vidal (2013) , como «el estudio de los principios que regulan el uso del lenguaje en la comunicación, es decir, las condiciones que determinan el empleo de enunciados concretos emitidos por hablantes concretos en situaciones comunicativas concretas, y su interpretación por parte de los destinatarios». Esto se debe a que esta disciplina considera factores extralingüísticos determinantes del uso del lenguaje que no tienen en cuenta otras (como la sintaxis) que se basan en un punto de vista puramente gramatical como: participante (emisor y destinatario), intención comunicativa, contexto verbal o conocimiento del mundo.

La conversación humana es un tipo de discurso que consiste en un diálogo oral o escrito no planificado entre dos o más personas que intervienen en turnos alternos para expresar sus ideas, para lo cual emplean lenguaje verbal y no verbal. Esta puede articularse en torno a uno o varios temas, que pueden variar fácilmente sin atenerse a unas pautas previamente establecidas. Además, está condicionada por el contexto, entendido como el conjunto de suposiciones que permiten la comprensión de un enunciado. Para poder representar adecuadamente el contexto de una conversación, se pueden utilizar dos teorías ampliamente usadas en el campo de la Inteligencia Artificial para describir

formalmente las situaciones en las que se lleven a cabo intercambios conversacionales: la teoría de marcos y la teoría de guiones.

4.3.1 Teoría de marcos de Minsky

Marvin Minsky, considerado uno de los padres de la Inteligencia Artificial, fue un científico estadounidense que elaboró lo que hoy se conoce como Teoría de marcos para explicar cómo actúa el ser humano ante situaciones nuevas para conseguir resolverlas con éxito. Lo que defiende la teoría es lo siguiente:

Cuando uno se encuentra con una situación nueva (o realiza un cambio sustancial en la visión del problema actual), se selecciona desde la memoria una estructura llamada marco. Este es un marco recordado que debe adaptarse para que se ajuste a la realidad, cambiando los detalles que sean necesarios.¹¹ (Minsky, 1974).

Un marco es una estructura de datos que representa una situación tipo, como entrar a un restaurante o subirse a un tren, a la que se le asocian diferentes tipos de información: cómo usar el marco, qué se puede esperar, qué hacer si no se cumplen las expectativas... Están compuestos por una colección de atributos –conocidos como ranuras– con valores asociados (ver figura 8) y describen una entidad del mundo, por ejemplo en la figura 9 se muestra un marco para el animal león con las ranuras que a las que se asocian los valores que definen a este tipo de mamífero. Puede decirse que es una red de nodos y de relaciones. Los niveles superiores son fijos y representan cosas que siempre son verdad sobre una situación hipotética; los niveles inferiores tienen muchas ranuras que se rellenan por datos específicos de cada situación. Cada ranura puede especificar condiciones que sus valores asignados deben cumplir. Podemos distinguir dos tipos de condiciones: las simples, especificadas por marcadores que pueden requerir la asignación de una ranura a una persona, un objeto u otro marco menor; y las complejas, que pueden especificar relaciones entre las cosas asignadas a diferentes ranuras.

¹¹ Traducción propia del original: «When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a structure called a Frame. This is a remembered framework to be adapted to fit reality by changing details as necessary.».

NOMBRE DEL MARCO	
NOMBRE DE LA RANURA	(NOMBRE DEL ATRIBUTO (VALOR))
NOMBRE DE LA RANURA	(NOMBRE DEL ATRIBUTO (VALOR))
NOMBRE DE LA RANURA	(NOMBRE DEL ATRIBUTO (VALOR))
(...)	(...)

Figura 8. Estructura general de un marco

LEÓN,NA	
ES-UN	(MAMÍFERO)
SEXO	(MASCULINO, FEMENINO)
EDAD	(INTERGER)
HÁBITAT	(SABANA)
DIETA	(CARNÍVORA)
Nº DE PATAS	(DEFECTO = 4)

Figura 9. Marco específico sobre el animal león

Las agrupaciones de marcos que guardan relación entre sí se unen en sistemas de marcos, conectados por el hecho de que el valor de un atributo de un marco puede ser otro marco (Rich y Knight, 1994). Tomando el ejemplo del marco de la figura 9, un sistema de marcos podría estar configurado por los marcos de león, mamífero y ser vivo (cada uno de ellos englobando al anterior, pues mamífero sería un valor del marco de ser vivo y, a su vez, león es un valor del de mamífero).

Los marcos de Minsky se utilizan para representar entidades del mundo real y no eventos o sucesiones de eventos. Para representar flujos de eventos se puede utilizar la teoría de guiones de Schank y Abelson

4.3.2 Teoría de los guiones de Schank y Abelson

Los orígenes de los guiones se remontan a la teoría de la dependencia conceptual desarrollada por Schank en 1975, pues este investigador siempre se centró en la representación del conocimiento y, en concreto, en el contexto del entendimiento del lenguaje. La idea clave sobre la que se fundamenta esta teoría es que todas las conceptualizaciones de eventos pueden representarse mediante un pequeño número de acciones realizadas por un actor o un objeto.

En base a este marco de trabajo, Schank y Abelson introdujeron en su libro *Scripts Plans Goals and Understanding* (1977), los conceptos de guiones, metas y temas para poder manejar el entendimiento de las situaciones cotidianas a las que se enfrenta el ser humano. Para este proyecto, el concepto más importante es el de guion, que surge de la necesidad de responder a las siguientes preguntas:

¿Cómo organizamos los seres humanos todo el conocimiento necesario para la comprensión? ¿Cómo sabemos cuál es la conducta apropiada en una determinada situación? Para ser más concretos, ¿cómo sabemos que, en un restaurante, la camarera nos traerá la comida que pedimos mientras que, si le pedimos un par de zapatos, o le pedimos comida cuando nos la encontramos en un bus, reaccionará como si hubiéramos hecho algo extraño? (Schank y Abelson, 1977, p. 37)

En su libro *Scripts Plans Goals and Understanding* (1977), dan respuesta a estas cuestiones mediante el concepto de guion. Para ello desarrollan el guion que seguimos en un restaurante, desde que entramos y buscamos sitio para sentarnos hasta que nos vamos, pagando o no; es decir, explican nuestro comportamiento y nuestra manera de actuar a través de esta idea.

Atendiendo a Rich y Knight (1994), a diferencia de los marcos, un guion es una estructura de ranura y relleno fuerte que describe una secuencia estereotipada de eventos en un contexto concreto. Proporciona un método de representación del conocimiento para secuencias comunes de eventos, por lo que permite construir una única interpretación coherente a partir de una colección de observaciones. De esta manera, puede considerarse que un guion es una gran cadena casual y que, por tanto, simplifica la visión sobre cómo se relacionan entre sí los eventos. De manera que un guion está formado por un conjunto de ranuras, a cada una de las cuales se le puede asociar información sobre el tipo de valores que puede contener. Los componentes más importantes son (Rich y Knight, 1994, p. 311):

- **CONDICIONES DE ENTRADA O PRECONDICIONES:** aquellas que deben satisfacerse antes de que puedan ocurrir los eventos descritos en el guion.
- **RESULTADOS:** las condiciones que deben ser ciertas tras suceder los eventos.
- **PROPS:** las ranuras que representan a los objetos involucrados.
- **ROLES:** las ranuras que hacen referencia a los participantes.
- **TRACK:** las variaciones específicas que se producen sobre un patrón representado por un guion.

- **ESCENAS:** las secuencias de eventos que tienen lugar.

Para ilustrar mejor estos conceptos, se incluye en las siguientes figuras el guion del restaurante propuesto por Schank y Abelson (Schank y Abelson, 1977, pp. 43–44):

Script: RESTAURANT

Track: Coffee Shop

Props: Tables

Menu

F-Food

Check

Money

Roles: S-Customer

W-Waiter

C-Cook

M-Cashier

O-Owner

Entry conditions: S is hungry.
S has money.

Results: S has less money
O has more money
S is not hungry
S is pleased (optional)

Scene 1: Entering

S **PTRANS** S into restaurant

S **ATTEND** eyes to tables

S **MBUILD** where to sit

S **PTRANS** S to table

S **MOVE** S to sitting position

Scene 2: Ordering

(menu on table)

(W brings menu)

(S asks for menu)

S **PTRANS** menu to S

S **MTRANS** signal to W

W **PTRANS** W to table

S **MTRANS** 'need menu' to W

W **PTRANS** W to menu

W **PTRANS** W to table

W **ATRANS** menu to S

S **MTRANS** food list to CP(S)

• S **MBUILD** choice of F

S **MTRANS** signal to W

W **PTRANS** W to table

S **MTRANS** 'I want F' to W

W **PTRANS** W to C

W **MTRANS** (**ATRANS** F) to C

C **MTRANS** 'no F' to W

W **PTRANS** W to S

W **MTRANS** 'no F' to S

(go back to *) or

(go to Scene 4 at no pay path)

C **DO** (prepare F script)

to Scene 3

Figura 10. Guion del restaurante parte I

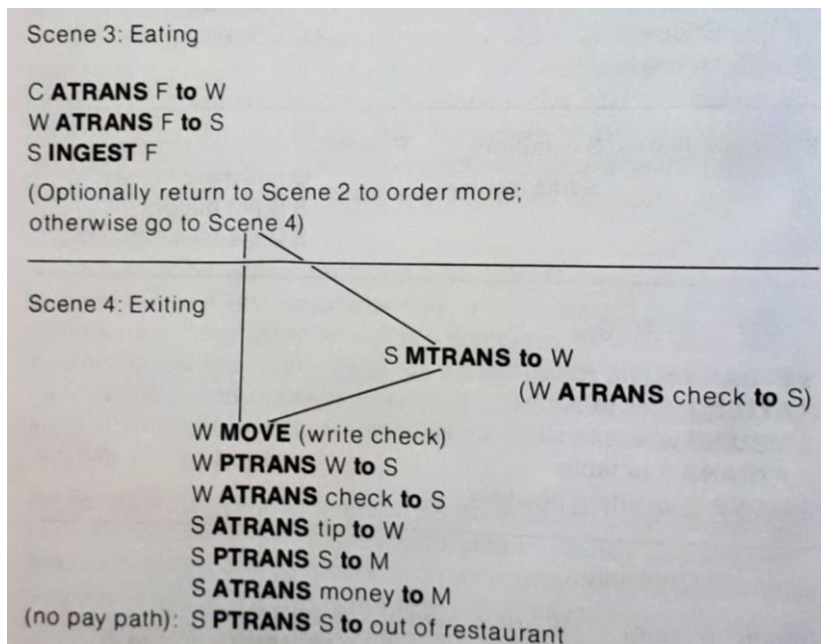


Figura 11. Guion del restaurante parte II

El uso de guiones es útil para representar eventos relacionados por causalidad (Rich y Knight, 1994). De esta forma, si se tiene la certeza de que un guion es apropiado para una situación dada, es posible predecir la ocurrencia de eventos utilizando estas relaciones

En este proyecto se ha utilizado este formalismo de guiones para estructurar los sistemas de representación de flujos conversacionales habitualmente utilizados hasta ahora.

4.4 Sistemas para representar flujos conversacionales

Para este apartado se ha realizado una búsqueda de biografía que se centre en este aspecto –la representación de los flujos de diálogo entre el usuario y los agentes– pero no se ha encontrado nada que hable de ello de forma directa. Si bien es cierto que hay algunos artículos académicos que hablan de la manera de representar la gestión de las tareas en los sistemas de diálogo en un nivel más cercano a la programación de la arquitectura del agente que a la programación de la conversación en el agente.

En la investigación realizada, se ha observado que los esfuerzos se han centrado en encontrar la mejor forma de representar y gestionar las tareas en los agentes conversacionales, como es el caso de González Bermúdez y Gatiús Vila (2010), quienes proponen en su artículo una arquitectura de sistemas de diálogo y un funcionamiento del sistema de diálogo muy visual que presentan la información de forma más clara mediante

esquemas, entendidos como una «representación gráfica o simbólica de cosas materiales o inmateriales» (Real Academia Española, 2014).

No obstante, la representación de los modelos de diálogo no ha recibido la misma atención que la representación de la arquitectura de los sistemas de diálogo y por ello, no se ha podido establecer una base objetiva y fundamentada para estas propuestas. Más bien, se han encontrado representaciones de flujos de conversación de manera indirecta en anexos de trabajos académicos (Duijst, 2017) donde se utiliza un flujograma (ver figura 12), o en la documentación de algunas herramientas de desarrollo de agentes conversacionales como la de Alexa de Amazon (Amazon), donde se emplea un guion al estilo tradicional literario (ver figura 13).

Encontrar un buen sistema de representación del flujo conversacional (representativo y usable) en estos casos tendría un impacto directo en la eficacia de los proyectos en las empresas ya que, facilitaría tanto su creación como su comprensión por parte de los distintos integrantes del equipo, lo que revierte en agilizar el proyecto acortando los tiempos de las tareas.

En este trabajo se utilizan ambos sistemas de representación, pero se modifican para tener un nivel mayor de estructura de la información aplicando la teoría de guiones anteriormente presentada.

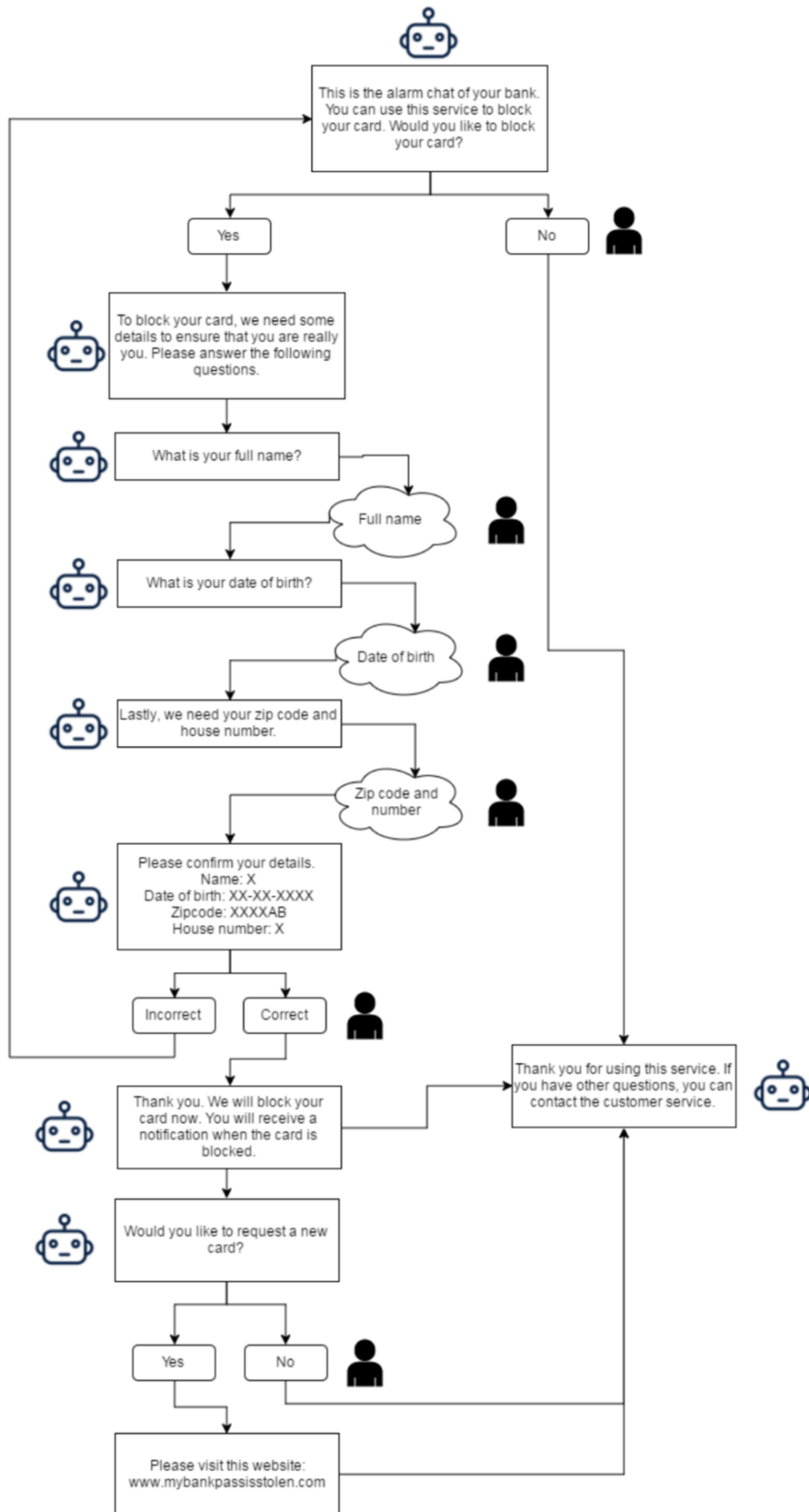


Figura 12. Flujograma representando una conversación

User: Alexa, open Paragon Investments.

Alexa: Welcome to Paragon. You can get a portfolio update or a security quote. Which one would you like?

User: Portfolio update.

Alexa: Your portfolio grew by 0.27% today. A total of \$322. Would you like a breakdown of the changes?

User: Yes.

Alexa: Crystal Cleaner grew by 0.35% for a total gain of \$165.03 (half second pause) Pizza Barn is up by 0.21% for a total gain of \$157. (half second pause) And Belle's Boats is up 1.41% for a gain of \$27.68. (half second pause) Do you want any other portfolio details?

User: That's it.

Alexa: What else would you like help with today?

User: Nothing, thanks.

Alexa: Okay, have a nice day.

User: Alexa, ask Paragon how my investments did today.

Alexa: Your portfolio grew by 0.27% today. A total of \$322. Would you like a breakdown of your investments?

User: Yeah.

Alexa: Crystal Cleaner grew by 0.35% for a total gain of \$165.03 (half second pause) Pizza Barn is up by 0.21% for a total gain of \$157. (half second pause) And Belle's Boats is up 1.41% for a gain of \$27.68. (half second pause) What other portfolio details would you like?

User: Exit.

Alexa: Okay, have a nice day.

Figura 14. Guion representando una conversación

4.5 Agentes conversacionales en el mundo empresarial

Como ya se ha visto en los apartados anteriores, la historia de los agentes conversacionales se remonta a los años sesenta con la invención de ELIZA. A lo largo de la segunda mitad del siglo XX y principios del siglo XXI, se han ido desarrollando diferentes agentes con diferentes técnicas y diversos objetivos. No obstante, a pesar de su larga historia, no ha sido hasta hace pocos años que han empezado ganar el interés de las empresas. Esto se ha debido principalmente a un cambio que se ha producido en cómo nos comunicamos las personas: en el 2016, 6.100.000.000 de personas utilizaban un teléfono con posibilidades de enviar y recibir SMS y las aplicaciones de mensajería instantánea eran utilizadas por más de 2.100.000 millones de personas (Dale, 2016, p. 815). Es decir, han aparecido plataformas gratuitas de comunicación instantánea (Facebook Messenger, WhatsApp, Telegram, Line, etc), accesibles por la gran mayoría de la población, que se han convertido en lugares idóneos para la publicación de agentes

Figura 13. Flujograma representando una conversación

conversacionales pues ofrecen a las empresas un canal para comunicarse con sus clientes de manera más directa. Si sumamos esto a las múltiples ventajas que ofrecen – atención a todas horas ininterrumpidamente, reducción de costes ya que reducen la carga de

peticiones que reciben los servicios de atención al cliente, mejoran la comunicación con el cliente, entre otros (Aracena, 2018)– se observa que la tendencia es que aumente el número de demandas de agentes por parte de las empresas.

En el siglo XXI podemos distinguir dos años muy importantes para los agentes conversacionales:

1. **2010.** En este año apareció por primera vez el agente de Apple, *Siri*. Hasta el momento los agentes conversacionales no habían despertado mucho interés en las grandes empresas tecnológicas y, a partir de ese momento muchos de los gigantes de la tecnología desarrollaron el suyo propio: en 2012 Google presentó a *Google Now* –que en 2016 sería sustituido por *Google Assistant*–, en 2014 Amazon, a *Alexa*, en 2015 Microsoft, a *Cortana* y en 2017 Samsung, a *Bixby*.
2. **2016.** Como bien indica Dale (2016) en el incipit de su artículo, 2016 fue el año de los chatbots –en el mundo angloparlante, por supuesto–. Esto se debe a dos factores principalmente: por un lado, la imposición de las plataformas de mensajería instantánea como medio de comunicación principal para gran parte de la sociedad mundial (sobre todo, de las generaciones más jóvenes); por otro lado, el gran ahorro de gastos que supone para las empresas implantar estas tecnologías en sus servicios.

Dada la importancia que está cobrando esta tecnología en los últimos años, las grandes compañías tecnológicas no han tardado en desarrollar diferentes plataformas que ponen a disposición de sus clientes todos los medios necesarios para que puedan desarrollar sus propios agentes conversacionales a través de una interfaz amigable y fácil de manejar. Por ello, actualmente, podemos encontrar a nuestra disposición una gran cantidad de plataformas para desarrollar agentes, siendo las más destacadas las siguientes: IBM Watson, LUIS de Microsoft, DialogFlow de Google, Chatfuel y Pandorabots.

A continuación, vamos a centrarnos en describir la herramienta presentada por Oracle en 2017 para la construcción de agentes, pues uno de los requisitos del cliente que contrató a Everis fue que se empleara esta herramienta para el desarrollo del agente conversacional cuyos modelos de diálogo se están desarrollando en este proyecto.

El servicio para crear agentes conversacionales de Oracle presenta un rasgo distintivo frente al resto de plataformas: la posibilidad de crear un único agente que sea multilingüe, es decir, permite crear un agente que, siempre y cuando sea desarrollado en inglés, pueda hablar en cualquier idioma mediante los paquetes de traducción automática que pueden integrarse (Google Translate o Traductor de Microsoft). De esta forma se

consigue que el agente pueda traducir de forma automática tanto los mensajes que recibe de los usuarios como las respuestas que debe dar.

Para desarrollar estos agentes, Oracle estructura su herramienta en varios componentes, siendo los siguientes los más importantes para la parte del desarrollo de los flujos conversacionales, es decir, de las conversaciones que va a mantener el agente con el usuario:

1. **INTENCIONES:** son categorías de acciones en las que se clasifican las frases del usuario según lo que necesita del agente; permiten que entienda lo que el usuario dice. Por ejemplo, las oraciones del tipo: «¿Qué es un libro?», «Quiero saber qué es un libro» o «definición de libro» se incluirían dentro de una intención denominada «Definición», pues lo que espera el usuario que haga el agente es que le devuelva la definición de «libro».
2. **ENTIDADES:** son palabras clave que añaden facilitan que el agente lleve a cabo la petición del usuario, pues permiten especificar más en ciertos aspectos de la intención al enriquecerla. Por ejemplo, si tenemos un agente para pedir pizzas a domicilio, habrá una intención que sea «Pedir Pizza», pero el precio de la pizza variará, naturalmente, en función de la pizza seleccionada. Para conseguir que el bot pueda devolver un precio diferente para cada tamaño de pizza, se le puede añadir una entidad a la intención que especifique que hay diferentes valores para la pizza. Es decir, añadimos una entidad “Tamaño” en donde añadimos los valores pequeño, mediano y grande. Con esto se consigue que, una vez programado, el agente pueda distinguir entre «Quiero pedir una pizza pequeña» y «Quiero pedir una pizza grande».
3. **OBOTML:** es la implementación propia de Yaml¹² realizada por el servicio de Oracle. Es el código mediante el cual se programa al agente.
4. **RESOURCE BUNDLE:** permite introducir las respuestas que debe dar el agente para cada intención. Este elemento tiene la ventaja de que permite introducir la misma respuesta en diferentes idiomas para que el agente conversacional pueda hablar con usuarios de diferentes países.
5. **Q&A¹³:** es un marco de trabajo especializado en preguntas frecuentes, no permite que haya ningún tipo de flujo conversacional pues solo se limita a responder a preguntas.

¹² Yaml (YAML Ain't Markup Language, en castellano), YAML no es un lenguaje de marcado es un formato para guardar objetos de datos con estructura de árbol que facilita la visualización de los datos ya que los presenta de forma muy legible.

¹³ Questions & Answers.

Es muy útil para los casos en los que las empresas quieren un agente para que responda a los usuarios las preguntas frecuentes.

5. Proyecto

En este capítulo se va a dar respuesta a la cuestión inicial del TFM: ¿cómo generar un guion para un agente conversacional en un dominio de conocimiento determinado?

Para ello, primero vamos a introducir el proyecto de Everis de forma global a través de los requisitos que debe cumplir. Segundo, detallaremos la metodología del proyecto que desarrolla este TFM, la creación de uno o varios modelos diálogo que representen de forma adecuada los diálogos que mantiene el agente con los usuarios dentro del dominio de la búsqueda de cursos online. Tercero, se expondrá quién es el agente y cuáles son sus objetivos. Cuarto, pasaremos a ver qué modelos se han diseñado para representar el mismo diálogo y por qué; para ello, especificaremos en qué consiste el flujo conversacional para el que se van a diseñar los modelos conversacionales y veremos los que se han generado en base a la información recopilada durante el estudio del estado de la cuestión.

5.1 Proyecto de Everis

En este apartado vamos a centrarnos en explicar por un lado el proyecto de Everis y qué objetivos persigue; y, por otro lado, el proyecto de este TFM con las diferentes propuestas planteadas y la metodología empleada para llevarlo a cabo.

5.1.1 Objetivos de Everis

El objetivo que se propone Everis para poder satisfacer la petición del cliente es desarrollar agente virtual que combine tecnología Big Data con una interfaz de interacción con el usuario, definida como un agente conversacional. Con esto se busca conseguir un agente que pueda ayudar a jóvenes en situación de desempleo a orientar su carrera laboral en el mundo digital.

La solución propuesta se llama Lía y es una orientadora educativa en el mundo digital, es decir, su función principal es ofrecer cursos y eventos que sean de utilidad a personas que buscan formarse en conocimientos específicos del mundo digital (HTML, SEO, CSS...). Esta formación se puede buscar en base a dos parámetros diferentes: carreras profesionales recomendadas o destrezas técnicas y sociales. Para ello, primero es necesario que el usuario se cree una cuenta y complete su perfil, aportando una serie de

datos que luego el agente conversacional de Everis utilizará para ver qué carreras profesionales se ajustan más al perfil que posee.

Así pues, el objetivo primordial es ser un punto referente de búsqueda de cursos formativos en conocimientos digitales, pero en ofreciendo una interfaz cercana que genere la idea en el usuario de estar hablando con un igual, con una persona, y no con un software programado. Para ello se ha desarrollado un software que accede a una base de datos en la que se alojan multitud de cursos procedentes de las principales plataformas de formación online (Udacity, Udemy, Coursera...) de los cuales se extrae la información clave que permite asignarlos a las destrezas técnicas correspondientes y mostrárselos al usuario.

El usuario puede acceder a los cursos de diversas maneras tales como la elección de una carrera profesional, de un tipo de destreza a mejorar o de una destreza en concreto directamente.

Para este tipo de proyecto se decidió emplear una metodología ágil conocida como Scrum pues las metodologías ágiles «son aquellas que permiten adaptar la forma de trabajo a las condiciones del proyecto, consiguiendo flexibilidad e inmediatez en la respuesta para amoldar el proyecto y su desarrollo a las circunstancias específicas del entorno» (Rosselló Villán, 2018). Esto tiene una clara ventaja que es que permite ir haciendo cambios y mejoras a lo largo del proyecto conforme se va desarrollando.

5.2 Proyecto del TFM

Tal y como se ha comentado en el capítulo de los objetivos de este trabajo, el proyecto de TFM consiste en desarrollar un modelo de diálogo que represente de forma adecuada los flujos conversacionales entre el agente de Everis y los usuarios en el dominio de conocimiento de formación en la red. Dado que no se encontraron en la empresa metodologías para el desarrollo sistemático de guiones de diálogo, una de las primeras tareas desarrolladas en este proyecto fue el diseño de una metodología nueva.

Antes de explicar qué metodología se decidió implantar para este proyecto y para los futuros, es necesario conocer más en profundidad lo que se produce durante la primera iteración, en la que se conceptualiza con el cliente el cómo va a ser el agente de Everis mediante una serie de sesiones denominadas talleres. En ellos, se realizan diversas técnicas para definir la personalidad del agente, ver quién es el público objetivo, entender cómo imagina el cliente que va a ser el producto final, etc., a fin de poder comenzar a desarrollarlo estando alineados con los intereses y expectativas que afloran.

Estos talleres suelen ser organizados por una persona del equipo de User Experience (UX) y de ellos sale un informe de resultados en el que se hace un resumen del contenido de cada sesión y de la información obtenida. Uno de los aspectos que se definieron en los talleres consistía en los diferentes casos de uso que debía cubrir el agente, cada uno de ellos desgranado con las diferentes opciones que el cliente veía necesario que se incluyera. Este ejercicio fue plasmado por la persona de UX del equipo en árboles de decisión a alto nivel, es decir, sin entrar en mucho detalle. Por lo tanto, era necesario profundizar más en cada uno de ellos para que luego el equipo de desarrollo no tuviera ninguna duda sobre cómo debía fluir cada una de las conversaciones.

Tras haber realizado una breve explicación del trabajo previo que se realiza en la definición de un agente conversacional, a continuación, se especifica la metodología planteada en este TFM para desarrollar modelos conversacionales:

1. Documentarse lo máximo posible a través de los informes de resultados de los talleres para adquirir todo el contexto del proyecto.
2. Extraer la información clave necesaria para dar forma a los diálogos: público objetivo y personalidad del agente. Esto es muy importante ya que a partir de esto se decide qué tipo de lenguaje empleará el agente y cómo se presentará al usuario: cercano o distante, humano o robot, coloquial o formal, etc.
3. Conceptualizar los árboles a bajo nivel, desgranando cada uno de los diagramas realizados por la persona de UX. En este caso se ha utilizado la herramienta gratuita draw.io. Para ello es esencial asignar formas y colores determinados a todos los posibles elementos del flujo, que en este caso serían: frases del usuario, respuestas del agente, botones, *webviews*¹⁴, gráficos, carruseles de vídeos o cursos, entre otros. Esto es de gran importancia ya que sirve para establecer una leyenda que se mantiene en todos los flujos pintados y así se consigue establecer un estándar que facilite el entendimiento en cualquier proyecto.
4. Revisar los flujos, antes de darlos por finalizados, si es posible por otro compañero de UX del equipo del agente para asegurarse de que no falta nada y de que la conversación parece, en efecto, natural.

El proceso para crear el flujo en sí es bastante sencillo, pues consiste en pensar cómo se desarrollaría una conversación de forma natural siguiendo unas pautas establecidas por el cliente; la dificultad adicional viene planteada por el hecho de que hay

¹⁴ Páginas web externas al agente que se abren pulsando un botón

que considerar qué es más cómodo para el futuro usuario dentro de la conversación: escribir todo o poder utilizar botones. Para mejorar la experiencia del usuario se estableció que el agente iba a dar respuestas aleatorias para generar una sensación más próxima a una conversación humana, por lo que, para no recargar en exceso el esquema del flujo, las respuestas se definen en un documento de Excel aparte junto con las respuestas que debe de dar el agente en caso de que se produzca un error o no se dé el caso idóneo.

Por supuesto, estos diagramas están sujetos a modificaciones conforme va avanzando el proyecto porque siempre son susceptibles de mejora, pero con esta metodología lo que se ha conseguido ha sido reducir los cambios a posteriori al mínimo, por lo tanto, se incrementa la efectividad del trabajo realizado.

5.3 Propuesta de modelo de conversación para el agente

A continuación, se va a proceder al modelado de la conversación que el agente de Everis, llamado Lía, va a mantener con los usuarios. Primero, se van a detallar los requisitos que debe cumplir el modelo mediante casos de uso (el conjunto de actividades que se realizan para llevar a cabo un proceso). Segundo, se hará una justificación de la elección los modelos que representan la misma conversación. Finalmente, se presentarán ambas propuestas de modelos de representación de la conversación.

5.3.1 Requisitos del modelo de conversación

El modelo de conversación se tiene que ajustar a una serie de requisitos que dependen del dominio de conocimiento del agente Lía. Estos requisitos se van a expresar en forma de los siguientes casos de uso:

- CASO DE USO 1. INICIAR SESIÓN O REGISTRARSE: cuando se abre el chat, el agente conversacional saluda de forma proactiva al usuario y le indica que inicie sesión o que se registre para poder continuar.

Si el usuario inicia sesión, se entiende que ya tiene su perfil completado por lo que Lía le lanza una pregunta que espera una respuesta abierta, por lo que el usuario tiene total libertad a la hora de responder. El agente estará programado de tal forma que pueda entender cualquiera de las intenciones para las que está entrenado; si el usuario dice algo para lo que no ha sido entrenado, le contestará con una de las respuestas definidas para el caso “no te entiendo”.

Si el usuario no tiene cuenta, entonces debe registrarse. Tras hacer el registro, el usuario debe completar su perfil rellenando los datos que se le piden en las pantallas modales que se abren (datos personales, estudios, destrezas sociales, intereses y

destrezas técnicas). Una vez completado el perfil, Lía le da un mensaje que sirve de tutorial de uso en el que le explica al usuario qué puede hacer y cómo puede emplear el agente conversacional.

- **CASO DE USO 2. SELECCIONAR CARRERA PROFESIONAL:** a continuación, el agente procede a recomendar al usuario las carreras profesionales con las que encaja su perfil según los datos que le ha facilitado. Estas son un máximo de seis, repartidas en dos carruseles con tres tarjetas cada uno. En cada tarjeta aparece una fotografía de la persona que habla de la carrera, un botón para ver el vídeo y otro para seleccionar la carrera que le interese.

Cuando el usuario seleccione una de las carreras profesionales, el agente le facilitará el número de ofertas de empleo que tiene esa carrera en su provincia y en España. Las ofertas para España vendrán representadas en un gráfico que muestre el ranking de las tres provincias donde hay más ofertas. Tras esto le ofrecerá las siguientes opciones:

- **Quiero formarme:** al seleccionar esta opción se pasa al siguiente caso de uso.
 - **Ver otras provincias:** el agente le pregunta al usuario la provincia sobre la que quiere obtener la información sobre el número de ofertas de empleo. Una vez hecho esto, el usuario puede preguntar por otra provincia o realizar una de las tres opciones restantes (quiero formarme, ver otras carreras e ir por destrezas).
 - **Ver otras carreras:** el agente le presenta la opción de ver las más demandadas en una provincia o en España, las más recomendadas (las mismas del inicio) o de ir por destrezas. Salvo en la última opción, le saca hasta dos carruseles de tres profesiones cada uno con los criterios aplicados por el usuario.
 - **Ir por destrezas:** se pasa a un caso de uso en el que se busca formación para una destreza técnica o una social.
- **CASO DE USO 3. SELECCIONAR UN CURSO:** Cuando el usuario decide formarse, el agente le enseña las diez destrezas más importantes que demanda la carrera escogida. Después, le enumera las que le faltan para perfeccionarla, es decir, para poder desempeñar adecuadamente las funciones relacionadas con ella en caso de que encontrara trabajo de eso.

El usuario elige la destreza para la que quiere buscar cursos, aunque siempre tiene la opción de decir que no le interesa ninguna. Una vez seleccionada, el agente le enseña hasta dos carruseles de tres cursos cada uno. Cada tarjeta de curso debe tener el logo

de la plataforma que oferta el curso y un botón de “ver detalle”, que abre una pantalla en la que se explica de forma muy visual la duración, una breve descripción, el nivel, el precio, quién lo imparte, la modalidad y el idioma.

Si no le interesa ninguno de los seis cursos ofertados, se le plantea la opción de buscar más cursos filtrando en base a tres criterios (precio, modalidad y nivel) o según las destrezas sociales o técnicas.

Por otro lado, además de estos casos de uso, se define también una serie de respuestas que aportan personalidad al agente como los saludos, las despedidas, los agradecimientos o los insultos, entre otros. Lía está planteada para que pueda guiar al usuario de forma más rápida y eficaz a través de una conversación natural y humana. Por tanto, está dotado de una personalidad definida con el cliente en los talleres previos al inicio del desarrollo. Así pues, encontramos que es una mujer de unos 32 años con un nivel educativo alto, cercana y con un carácter proactivo dispuesto a explicar cualquier cosa al usuario.

Se va a poder apreciar que este es un flujo muy dirigido en el que el usuario tiene pocas opciones de escritura libres, esto se debe a un motivo principal: este agente está planteado para que se utilice en dispositivos móviles esencialmente y está dirigido a un público joven. Tras realizar varias pruebas de usuario se observó que la tendencia general era que preferían botones a escribir desde el móvil ya que era más cómodo y no les resultaba tan pesado como tener que teclear todo.

5.3.2 Metamodelo del modelo de conversación

El metamodelo se refiere al mecanismo de representación del modelo o flujo de conversación. En el apartado 4.4 se presentó el problema de falta de bibliografía académica centrada en metamodelos de representación de diálogos. De hecho, los que se han presentado se encontraron en los anexos de artículos centrados en otros temas como la experiencia de usuario o en páginas web de documentación para el desarrollo de agentes conversacionales (ej. el caso de *Alexa* de Amazon). Basándonos en estos ejemplos, proponemos el uso de dos metamodelos para representar los flujos conversacionales: el flujograma y el guion.

- (1) El flujograma es el tipo de esquema que presenta Duijst (2017) en los anexos de su tesis. En este caso los mensajes del agente están indicados con una imagen de un robot al lado del recuadro de diálogo. Sin embargo, estos flujogramas están poco estructurados lo cual dificulta su comprensión y

utilización por parte de los informáticos que programarán los modelos de conversación. Para estructurar los flujogramas de Duijst (2017), proponemos aplicar la teoría de guiones de Schank y Abelson de la forma siguiente: cada caso de uso se puede representar como un guion compuesto por una acción principal y en donde hay diferentes ranuras que adquieren valores concretos en la conversación. La siguiente figura ilustra una de las opciones del caso de uso 1, en el que el usuario ya se ha registrado y quiere iniciar sesión y muestra, el patrón que siguen las escenas que tienen lugar.

Guion: Acceder a cuenta Track: iniciar sesión Objetos: Móvil Plataforma de mensajería Red social Roles: A = Agente U = Usuario	Escena 1: iniciar conversación con A U ATRANS la conversación con A en su móvil A MTRANS a U A ATRANS a U que inicie sesión o se registre U ATRANS iniciar sesión
	Escena 2: seleccionar opción de inicio de sesión A ATRANS a U que elija una opción para iniciar sesión U ATRANS una red social
Condiciones de entrada: U quiere iniciar sesión Resultados: U accede a su cuenta	Escena 3: acceder a cuenta U ATRANS a su cuenta

Figura 15. Ejemplo de guion para el caso de uso 1

- (2) La idea del guion “literario” ha sido tomada de la documentación de Amazon para la configuración de una *skill* para su agente llamado Alexa. Este es el esquema más clásico y el que se utiliza para representar la mayoría de las conversaciones. En este caso, de igual forma, se usará el concepto de guion con ranuras para estructurar el contenido de la conversación

5.3.3 Propuesta 1: modelo de conversación en forma de flujograma

En esta propuesta se representa la conversación con un diagrama de flujo de forma muy visual, como se puede observar en los anexos I, II, III y IV. Aunque la conversación es un continuo y podría representarse dentro de la misma imagen, se ha decidido separarla en diferentes imágenes, que se corresponden con cada uno de los casos de uso.

Con ese modelo se consigue representar de forma muy visual el transcurso de la conversación, pues cada una de las formas representa uno de los elementos del flujo: las cajas moradas, las respuestas del agente; las cajas amarillas, las respuestas del usuario;

las cajas azules, los botones; las cajas rojas junto con las azules, los carruseles; las naranjas, los gráficos; y las nubes amarillas, los momentos en los que se da un «campo abierto» (el usuario puede contestar libremente).

5.3.4 Propuesta 2: modelo de conversación en forma de guion “literario”

En esta propuesta se representa la conversación de forma tradicional, como si se tratara de un guion de una obra de teatro, como se aprecia en el anexo V. Las acotaciones que aportan algo de contexto o que hacen algún comentario que debe tenerse en cuenta se resaltan por ir entre corchetes y en negrita. Las diferentes opciones por las que puede continuar el usuario se resaltan con un subrayado. Se indican las intervenciones del agente conversacional mediante “Bot:” al inicio de cada frase suya.

Es menos visual, pero permite una descripción más detallada de algunas casuísticas concretas.

6. Validación y pruebas

Tras haber realizado las dos propuestas para representar el flujo conversacional, se va a proceder a realizar una encuesta a los desarrolladores principales del proyecto para validar la usabilidad de los dos modelos en términos de visualización y de comprensión.

Para ello, este capítulo se plantea dividido en tres apartados: uno en el que se presenta la encuesta elaborada para evaluar los puntos que se consideran clave a la hora de programar el agente partiendo de la conceptualización del flujo; otro, en el que se presentan los resultados; y un último apartado en el que se comentan los resultados obtenidos.

6.1 Encuesta

El principal objetivo que se busca con estos modelos de representación es encontrar la mejor forma de representar la conversación que el agente de Everis va a mantener con los usuarios para que los programadores sean lo más autónomos posibles a la hora de desarrollarlo, es decir, sin que tengan que estar pidiendo ayuda para su comprensión de manera constante.

La encuesta que se plantea en este apartado sirve para comprobar cuál de las dos propuestas, flujograma o guion, es la más usable a la hora de indicar a los programadores cómo tiene que discurrir la conversación entre el agente conversacional y el usuario. Por lo tanto, se han realizado las siguientes preguntas:

1. ¿Te parece útil que se represente el flujo conversacional antes de que lo desarrolles?
2. ¿Cuál de los dos modelos presenta de forma más clara la información?
3. ¿Qué modelo te parece más sencillo y fácil de entender?
4. ¿Sientes que necesitarías ayuda para poder pasar a código los modelos?
5. ¿Crees que le falta algo a alguno de los modelos?
6. En caso afirmativo, ¿por qué?
7. ¿Con qué modelo te quedarías?
8. ¿Por qué?

6.2 Resultados obtenidos

La encuesta se ha realizado a través de la plataforma Google Formularios ya que facilita mucho obtener los resultados y, además, da más flexibilidad a los encuestados de realizarla. Se envió a doce personas, la mayoría con perfiles informáticos desarrollando agentes; de entre ellas, algunas pertenecen a otros equipos en los que la metodología de

conceptualización de flujos conversacionales es diferente. A continuación, se presentan los resultados obtenidos en cada pregunta¹⁵:

1. ¿Te parece útil que se represente el flujo conversacional antes de que lo desarrolles?

13 respuestas

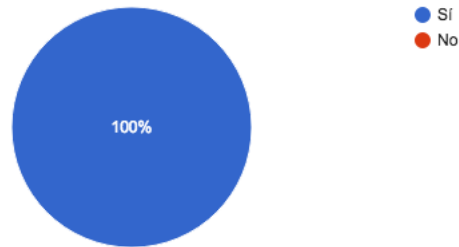


Figura 16. Gráfico de los resultados de la pregunta 1

2. ¿Cuál de los dos modelos presenta de forma más clara la información?

13 respuestas

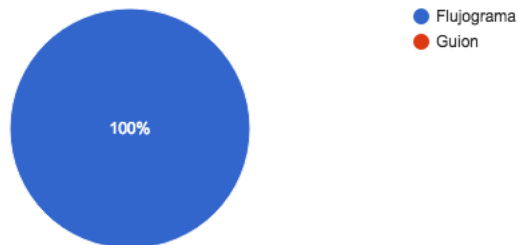


Figura 17. Gráfico de los resultados de la pregunta 2

3. ¿Qué modelo te parece más sencillo y fácil de entender?

13 respuestas

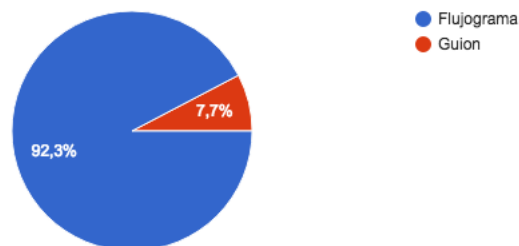


Figura 18. Gráfico de los resultados de la pregunta 3

4. ¿Sientes que necesitarías ayuda para poder pasar a código los modelos?

¹⁵ Las respuestas de las preguntas 6 y 8 se han corregido para evitar errores ortográficos. Los originales se pueden encontrar en los anexos VI y VII.

13 respuestas

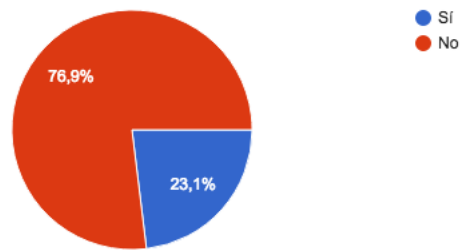


Figura 19. Gráfico de los resultados de la pregunta 4

5. ¿Crees que le falta algo a alguno de los modelos?

13 respuestas

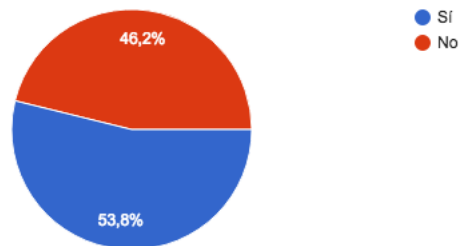


Figura 20. Gráfico de los resultados de la pregunta 5

6. En caso afirmativo, ¿por qué?

- Para el caso del guion creo que sería más cómodo contar con un guion por cada caso de uso o camino único que puede realizar el usuario, ya que en el formato actual es bastante complicado seguir los posibles flujos a implementar.
- Para pasar a código los modelos necesitaría más especificación de los atributos de cada elemento
- En el modelo del flujograma veo necesario añadir una leyenda que facilite al usuario el entendimiento del modelo sin necesidad de leerlo, solo viendo los coles. En caso del guion veo que es difícilmente comprensible porque:
 - Faltaría mejorar formato para facilitar que se entienda en un vistazo si es el agente el usuario
 - Igualmente, con desambiguaciones y comentarios del lingüista, al no estar distinguidos cuesta entender el flujo a la primera en una lectura rápida.
- Incluiría los mensajes que introduce el usuario y una leyenda en los diagramas que indique qué es cada elemento del diagrama.
- Flujograma:
 - ¿Qué continúa tras registrarme? ¿Campo abierto? ¿Carreras? No sé cómo se llega al flujo 2 "recomendación de carreras". No sé cómo se llega al flujo 3 ya

que empieza con "has escogido una carrera interesante", supongo que es tras el "quiero formarme".

- No se especifica que las carreras y los cursos deben llevar logos e imágenes ni que hacen los botones "ver video" ni "ver detalles" de cada uno (en el guion si se detalla esto).
- No se detalla con exactitud que ocurre tras seleccionar los filtros y mostrar el carrusel: ¿Campo abierto? ¿Puedo escoger más filtros? ¿Lía me muestra algún botón? (En el guion si se detalla esto).

Guion:

- El orden de las opciones desorienta a veces y es fácil perderse
 - Dividiría el guion según el flujo para entender de manera clara cuando estoy en el flujo de registro y cuando cambio al flujo de carreras, por ejemplo
- En caso de error que pasaría
 - Asignaría iconos (rombos, rectángulos, círculos) a cada interacción posible en cada punto, como una decisión, un inicio o final de conversación, etc.
 - Las varias respuestas que pueden ocurrir tras una transición de estado a estado, deberían de estar representado en el diagrama de flujo, en vez de explicar que ha de verse un Excel.

7. ¿Con qué modelo te quedarías?

13 respuestas

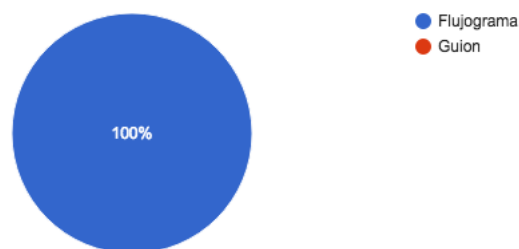


Figura 21. Gráfico de los resultados de la pregunta 7

8. ¿Por qué?

- Desde el punto de vista del desarrollador, contar con el flujograma ayuda a poder interpretar correctamente todas las casuísticas para tener en cuenta, teniendo una visión global del proceso a implementar y quedando claros todos los casos de uso.
- Es más claro, a primera vista se más rápido y es mucho más fácil de entender.
- Me parece mucho más visual, fácil de entender de un vistazo, y mucho más cercano al lenguaje que utilizan los equipos de desarrollo para su posterior implementación.

- Me parece mucho más fácil de entender visualmente.
- Creo que está mejor representada la información y que es una interacción mucho más real con una persona. En cuanto a pasar a código quizá sí que sea más difícil, pero creo que tiene más ventajas en cuanto a usabilidad que el guion.
- Es mucho más fácil de seguir, de visualizar y (sobre todo) aclara muchísimo más a la hora de programar. Gracias a ellos podría dividir mi código de manera eficaz en diferentes archivos de manera más ordenada y entendería que hace cada cosa. El guion es muchísimo más caótico y es muy fácil perderse u olvidar que todavía quedaba una opción más que había aparecido con anterioridad al listar los botones. Muchas veces también se pierde el hilo de si estamos en un flujo o si hemos salido de él, es decir, si las respuestas de Lía siguen siendo parte de la Opción X o si nos encontramos en un flujo aparte. De todas maneras escribir esto último de manera clara sería imposible ya que acabarían apareciendo cosas como "Opción 2.1.4.2". El flujograma aun así necesita de más detalles en algunos puntos y creo que nunca se debería terminar un diagrama sin saber a dónde va, es decir, los diagramas siempre deberían terminar en un "campo abierto", en un "Se va al caso de uso X" o en un "vuelve al punto X" (como hace el segundo diagrama con "quieres buscar otra provincia -> Sí -> 2") ya que el programador no puede dejar el agente "muerto", siempre debe quedarse en un punto donde el usuario decide cual es la siguiente opción (ejemplo: cuando el usuario elige "formarme" el programador no sabe que debería hacer Lía). De todas maneras, esto ocurre en ocasiones puntuales y son cuestiones sencillas de resolver, pero para ello es necesario contactar con el/la lingüista encargado/a (es por ello por lo que he marcado la opción afirmativa de necesitar ayuda al pasarlo a código).

En definitiva, el flujograma me ayudaría mucho más a la hora de programar a Lía y solo necesitaría ayuda externa para cuestiones y detalles más profundos.

- El flujograma me parece un modelo mucho más visual y fácil de entender a la hora de desarrollar las funcionalidades que se plantean.
- Porque tienes una visión global del flujo. Te permite diferenciar de un vistazo que partes se pueden desarrollar en paralelo y las que son parecidas para no implementar dos veces lo mismo. También te permite aislar zonas, es decir, al tener una vista en general del flujo puedes decidir de una manera rápida hasta donde puedes desarrollar o que partes hacer en el futuro sin dejar pobre la funcionalidad.

- El flujograma contiene más información visual para el programador y da lugar a menos confusión que el guion
- Queda más claro y de un vistazo puedes ver todo el caso de uso para poder tomar mejores decisiones técnicas. También añadiría más detalle para saber que debo mostrarle al usuario (texto, botones, imágenes, gif, etc.) y además algún modo de poder ver la URL/diseño de esos elementos gráficos.
- Es una vista esquemática que se entiende muy claramente y su traducción a código es bastante directa.
- Fácil comprensión y asimilación

6.3 Discusión de resultados

Tras haber analizado la encuesta se han extraído los siguientes resultados:

1. Todos los encuestados consideran que la representación del flujo conversacional en una etapa previa al desarrollo es útil.
2. Todos los encuestados creen que el flujograma representa la información de forma más clara.
3. Doce de trece encuestados consideran que el flujograma es más sencillo y fácil de entender.
4. Diez de trece encuestados no necesitarían ayuda para poder pasar a código el flujo conversacional plasmado en las dos propuestas.
5. Siete de trece encuestados creen que a los modelos les faltan algunos elementos tales como:
 - En el caso del flujograma: una leyenda, más conexión entre los diagramas de cada caso de uso para no perder el hilo, iconos que representen tipos de interacciones en los puntos de decisión.
 - En el caso del guion: segmentarlo en casos de uso, mejorar el formato para dividir de forma más clara cada elemento.
6. Los trece encuestados se quedarían con el flujograma por encima del guion ya que, en general, es más visual y más fácil de entender, pues aporta una visión global del flujo. Además, es más cercano a lo que utilizan los equipos de desarrollo para la implementación.

7. Resumen, conclusiones y trabajo futuro

En el presente trabajo se ha buscado obtener una respuesta a la cuestión inicial que fue planteada al inicio de este documento: ¿cuál es el modelo de conversación adecuado para el agente conversacional en el dominio de formación en línea?

Para responder, se ha realizado, en primer lugar, una revisión de los sistemas de diálogo, sistemas que buscan llevar a cabo lo más fielmente posible una conversación natural, tras realizar una introducción al Procesamiento del Lenguaje Natural y la representación del conocimiento, nos centramos teorías relacionadas con la comunicación humana dentro del campo de la Inteligencia Artificial, en concreto con lo relacionado con las teorías de Marvin Minsky y Schank y Abelson. Después, se hizo un recorrido global de los agentes conversacionales, deteniéndonos en los diferentes tipos existentes. Finalmente exploramos la situación actual de desarrollo de agentes conversacionales, sobre todo desde el punto de vista de las empresas, que se valen de numerosas herramientas y plataformas para crear sus agentes conversacionales.

Respecto a los sistemas para representar los flujos de conversación, se encontró que no existen trabajos académicos sobre la modelización de los flujos conversacionales. Los sistemas encontrados provienen, de forma indirecta, de ejemplos en trabajos centrados en la arquitectura y el funcionamiento de los agentes conversacionales o de la documentación de las herramientas software de construcción de agentes. Básicamente se trata de dos sistemas: del flujograma y el guion «literario».

Estos sistemas, sin embargo, no están suficientemente estructurados como para representar de forma precisa el modelo de conversación, al menos, de forma que puedan directamente ser utilizados por los desarrolladores informáticos para programar estas conversaciones. En este sentido, este TFM propone un sistema de estructuración basado en la teoría de guiones de Schank y Abelson.

Para poder valorar de forma objetiva cuál de los dos metamodelos propuestos es más usable para modelar los flujos de conversación y para servir de base a la programación del agente conversacional, se realizó una evaluación empírica basada en encuestas. Del análisis de resultados se obtuvo que el metamodelo de flujograma parece ser el más usable para los especialistas en informática, principalmente porque permite visualizar más rápida y fácilmente el algoritmo de interacción en lenguaje natural del agente conversacional con el usuario. No obstante, se detectaron algunas propuestas de mejora que deben ser consideradas para mejorar las futuras aplicaciones del modelo de

flujos conversacionales. Entre ellas, cabe destacar la idea de que cuanto más específico y detallado sea el modelo, mejor será la comprensión y autonomía de los programadores en la programación de dichos flujos.

Finalmente, conviene tener en cuenta que los resultados de este proyecto están limitado a un dominio de conocimiento concreto, por lo cual es recomendable que se contrasten, en el futuro, con otras experiencias sobre la representación formal de las conversaciones de agentes conversacionales. De esta manera se podrá obtener un modelo general para la representación precisa e interdisciplinar de cualquier flujo conversacional.

En esta línea de trabajo futuro, también debería considerarse el problema, no tratado en este TFM, de la detección e incorporación al modelo de conversación del contexto de dicha conversación como mecanismo para mejorar la eficacia de un agente conversacional.

8. Referencias

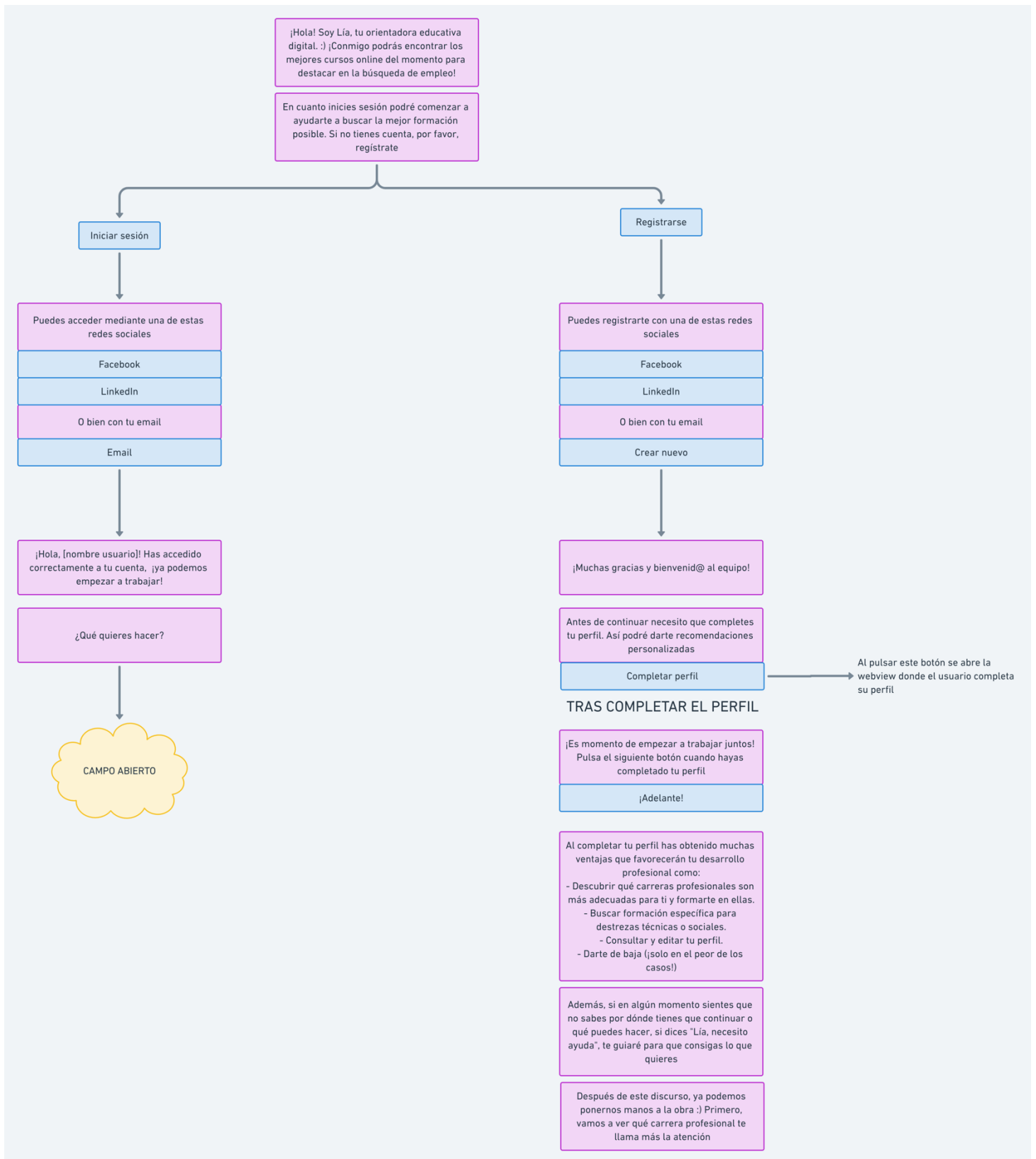
- Allen, J. (1995). *Natural language understanding* ([2nd ed.]). Redwood City, California: Benjamin/Cummings Publishing Company.
- Amazon. *Design Process / Amazon Alexa Voice Design Guide*. Recuperado de <https://developer.amazon.com/designing-for-voice/design-process/#write-scripts>
- Aracena, C. (2018). *5 razones para integrar un chatbot en tu negocio*. Recuperado de <https://chatbotchile.cl/blog/5-razones-para-integrar-un-chatbot/>
- Arthur, C. (2012). *The history of smartphones: timeline*. Recuperado de <https://www.theguardian.com/technology/2012/jan/24/smartphones-timeline>
- Barr, A. y Feigenbaum, E. A. (1981). *The handbook of artificial intelligence. Vol. I*. Reading, Mass.: Addison-Wesley.
- Blasco, L. (2017, 1 noviembre). ¿Puede Woebot, el bot parlante de Facebook, realmente ayudar a combatir la depresión? *BBC Mundo*. Recuperado de <https://www.bbc.com/mundo/noticias-41817941>
- Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., Thompson, H. y Winograd, T. (1977). GUS, a frame-driven dialog system. *Artificial Intelligence*, 8, 155–173. [https://doi.org/10.1016/0004-3702\(77\)90018-2](https://doi.org/10.1016/0004-3702(77)90018-2)
- Cortez Vásquez, A., Vega Huerta, H. y Paquinona Quispe, J. (2009). Procesamiento del lenguaje natural. *Revista de Ingeniería de Sistemas de Informática*, 6, 45–54. Recuperado de http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/risi/2009_n2/v6n2/a06v6n2.pdf
- Dahl, D. A. (2013). Natural Language Processing: Past, Present and Future. En A. Neustein y J. A. Markowitz (Eds.), *Mobile speech and advanced natural language solutions* (pp. 49–73). New York: Springer. https://doi.org/10.1007/978-1-4614-6018-3_4
- Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22, 811–817. <https://doi.org/10.1017/S1351324916000243>
- Duijst, D. (2017). Can we Improve the User Experience of Chatbots with Personalisation? (Tesis). University of Amsterdam. Recuperado de https://www.researchgate.net/publication/318404775_Can_we_Improve_the_User_Experience_of_Chatbots_with_Personalisation

- Escandell Vidal, M. V. (2013). *Introducción a la pragmática* (1ª ed. en esta presentación). *Ariel Letras*. Barcelona: Ariel.
- García Serrano, A. (2016). *Inteligencia artificial : fundamentos, práctica y aplicaciones* (2ª ed. rev). San Fernando de Henares, Madrid: RC Libros.
- González Bermúdez, M. y Gatiús Vila, M. (2010). Estructura y gestión de tareas en un sistema de diálogo para acceder a servicios web. *Procesamiento del lenguaje natural*, 45, 183–190. Recuperado de http://rua.ua.es/dspace/bitstream/10045/14727/1/PLN_45_25.pdf
- Griol, D., Zoraida, C., López-Cózar, R. y Gutiérrez, A. (2009). Utilización de los sistemas de diálogo hablado para el acceso a la información en diferentes dominios. En *Actas de la Segunda Conferencia Internacional sobre brecha digital e inclusión social*. Madrid. Recuperado de <http://hdl.handle.net/10016/12397>
- Hirschberg, J. y Manning, C. D. (2015). Advances in natural language processing. *Science (New York, N.Y.)*, 349, 261–266. <https://doi.org/10.1126/science.aaa8685>
- Isbell, C. L., Kearns, M., Singh, S., Shelton, C. R., Stone, P. y Kormann, D. (2006). Cobot in LambdaMOO: An Adaptive Social Statistics Agent. *Autonomous Agents and Multi-Agent Systems*, 13, 327–354. <https://doi.org/10.1007/s10458-006-0005-z>
- Jurafsky, D. y Martin, J. H. (2017). *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*. Recuperado de <https://web.stanford.edu/~jurafsky/slp3/>
- Khan, R. y Das, A. (2018). *Build Better Chatbots*. Berkeley, CA: Apress.
- Minsky, M. (1974). *A Framework for Representing Knowledge*. Cambridge, MA, USA.
- Portolés, J. (2004). *Pragmática para hispanistas. Letras universitarias*. Madrid: Síntesis.
- Psychology Today. (2018). *Person-Centered Therapy | Psychology Today*. Recuperado de <https://www.psychologytoday.com/us/therapy-types/person-centered-therapy>
- Real Academia Española. (2014). *Diccionario de la lengua española*. Recuperado de <http://dle.rae.es/?id=GgeQ9Uy>
- Rich, E. y Knight, K. (1994). *Inteligencia artificial* (2a. ed.). Madrid: McGraw-Hill.
- Rosselló Villán, V. (2018). *Las metodologías ágiles más utilizadas y sus ventajas dentro de la empresa*. Recuperado de <https://www.iebschool.com/blog/que-son-metodologias-agiles-agile-scrum/>

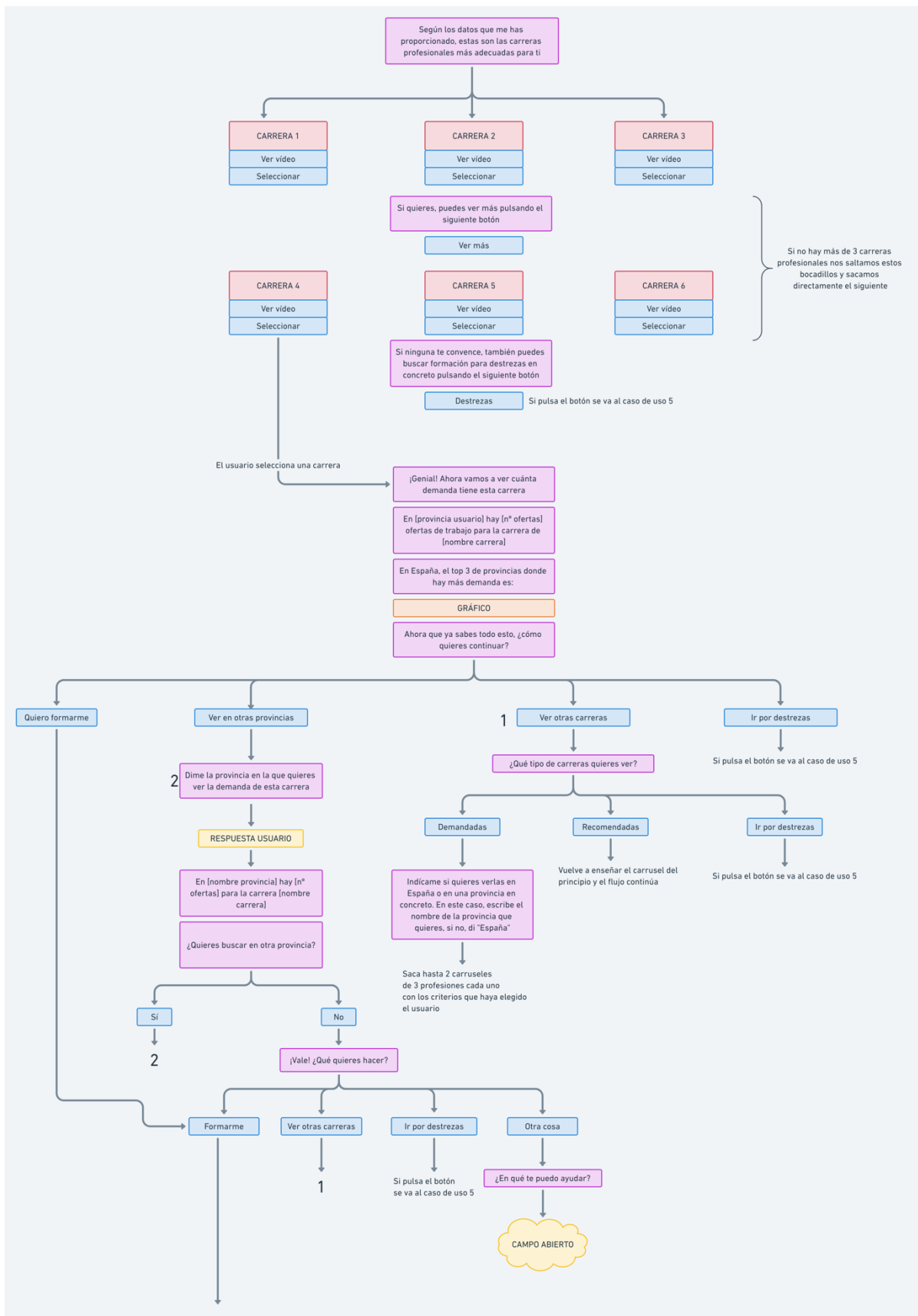
- Sapir, E. (1980). *Language: an introduction to the study of speech*. San Diego: Harcourt Brace Jovanovich (Trabajo original publicado en 1921).
- Schank, R. C. y Abelson, R. P. (1977). *Scripts, plans, goals and understanding: an inquiry into human knowledge structures. Artificial intelligence series*. Hillsdale: Lawrence Erlbaum Associates.
- Shang, L., Lu, Z. y Li, H. (2015). Neural Responding Machine for Short-Text Conversation. En C. Zong y M. Strube (Eds.), *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 1577–1586). Stroudsburg, PA, USA: Association for Computational Linguistics. <https://doi.org/10.3115/v1/P15-1152>
- Statista. (2018). *Most popular messaging apps 2018*. Recuperado de <https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/>
- Visual Capitalist. (2016). *Infographic: The Evolution of Instant Messaging*. Recuperado de <http://www.visualcapitalist.com/evolution-instant-messaging/>
- Weizenbaum, J. (1966). ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9, 36–45. <https://doi.org/10.1145/365153.365168>

9. Anexos

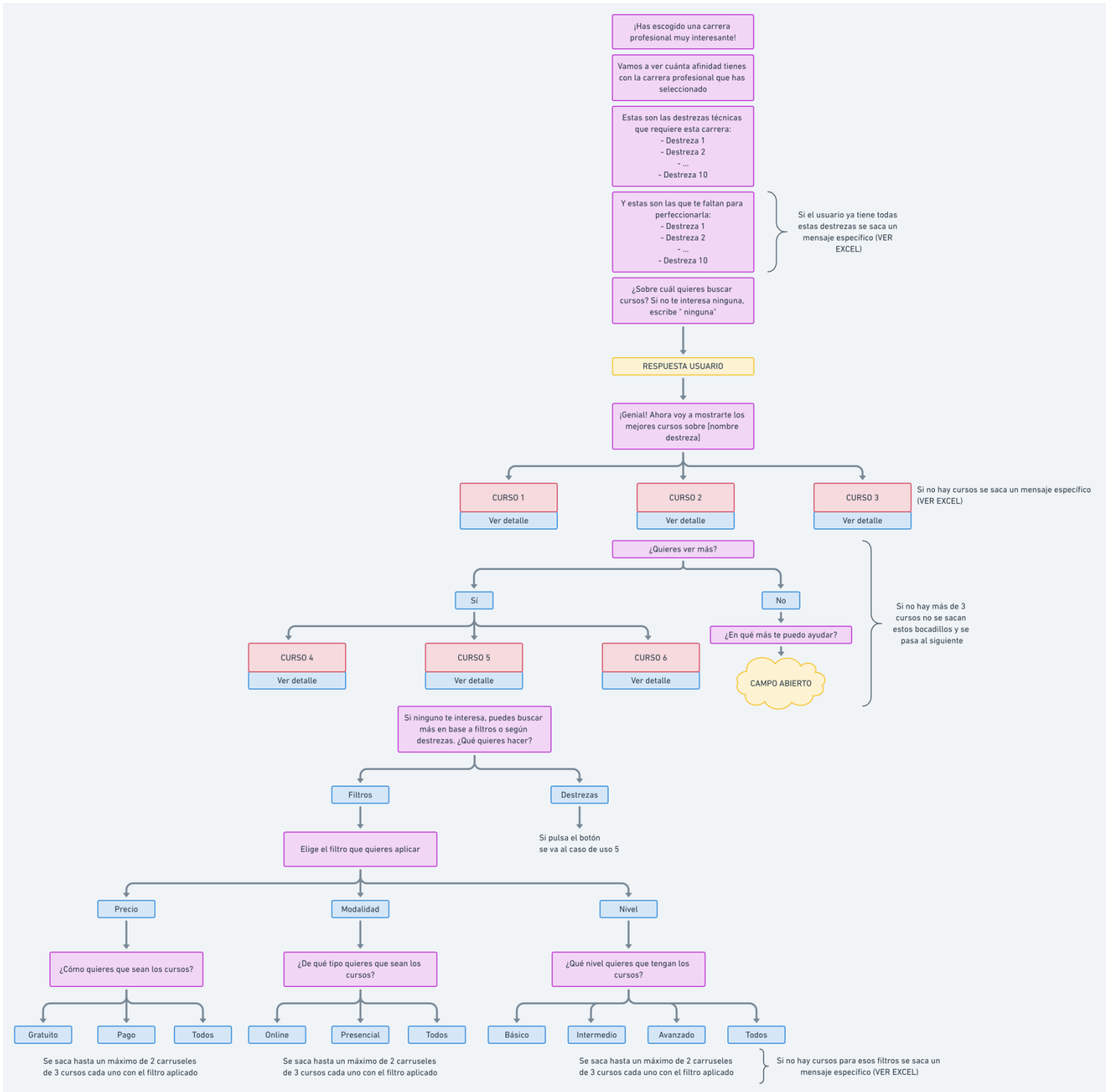
Anexo I. Caso de uso 1 – Registro y completar perfil



Anexo II. Caso de uso 2 – Seleccionar carrera profesional



Anexo III. Caso de uso 3 – Itinerario formativo



Anexo IV. Excel con mensajes para casos excepcionales en los flujogramas

Error general (no se puede iniciar sesión, no se puede completar el perfil..)	¡Vaya! Parece que hay un problema que no deja que las cosas funcionen como deberían... Inténtalo de nuevo más tarde, por favor
Mensaje para cuando el usuario tiene todos los conocimientos que requiere el perfil	¡Qué bien! Tienes todos los conocimientos para poder buscar ofertas de empleo para UX / UI
Mensaje para cuando no hay cursos	Ahora mismo no hay ningún curso ofertado para esta destreza. ¿Qué otra cosa quieres hacer?
Mensaje para cuando no hay cursos para ese filtro	Ahora mismo no hay ningún curso ofertado con las características que buscas. ¿En qué más te puedo ayudar?

Anexo V. Flujo representado en un guion “literario”

[CADA SALTO DE LÍNEA INDICA UN BOCADILLO NUEVO]

BOT: ¡Hola! Soy Lía, tu orientadora educativa digital. :) ¡Conmigo podrás encontrar los mejores cursos online del momento para destacar en la búsqueda de empleo!

En cuanto inicies sesión podré comenzar a ayudarte a buscar la mejor formación posible.

Si no tienes cuenta, por favor, regístrate

[Botones: iniciar sesión, registro]

OPCIÓN 1: USUARIO ELIGE “INICIAR SESIÓN”

B: Puedes acceder mediante una de estas redes sociales

[Botones: Facebook, LinkedIn]

O bien con tu email

[Botones: Email]

[Después de que el usuario seleccione una opción]

¡Hola, [nombre de usuario]! Has accedido correctamente a tu cuenta, ¡ya podemos empezar a trabajar!

¿Qué quieres hacer?

[Campo abierto, el usuario puede decir lo que quiera y el agente debe entenderlo]

OPCIÓN 2: USUARIO ELIGRE “REGISTRARSE”

B: Puedes registrarte con una de estas redes sociales

[Botones: Facebook, LinkedIn]

O bien con tu email

[Botones: Crear nuevo]

[Después de que el usuario seleccione una opción]

¡Muchas gracias y bienvenid@ al equipo!

Antes de continuar necesito que completes tu perfil. Así podré darte recomendaciones personalizadas

[Botones: Completar perfil]

[Tras completar el perfil]

¡Es momento de empezar a trabajar juntos! Pulsa el siguiente botón cuando hayas completado tu perfil

[Botones: ¡Adelante!]

Al completar tu perfil has obtenido muchas ventajas que favorecerán tu desarrollo profesional, como:

- Descubrir qué carreras profesionales son más adecuadas para ti y formarte en ellas.
- Buscar formación específica para destrezas técnicas o sociales.
- Consultar y editar tu perfil.
- Darte de baja (¡solo en el peor de los casos!)

Además, si en algún momento sientes que no sabes por dónde tienes que continuar o qué puedes hacer, si dices “Lía, necesito ayuda”, te guiaré para que consigas lo que quieres

Después de este discurso, ya podemos ponernos manos a la obra :) Primero, vamos a ver qué carrera profesional te llama más la atención

Según los datos que me has proporcionado, estas son las carreras profesionales más adecuadas para ti

[Carrusel de 3 tarjetas. Cada tarjeta debe tener: foto de la carrera profesional, botón de ver vídeo y de seleccionar]

Si quieres, puedes ver más pulsando el siguiente botón

[Botón: Ver más]

[Carrusel de 3 tarjetas. Cada tarjeta debe tener: foto de la carrera profesional, botón de ver vídeo y de seleccionar]

Si ninguna te convence, también puedes buscar formación para destrezas en concreto pulsando el siguiente botón

[Botón: destrezas. Si pulsa este botón se va al caso de uso 5 – Destrezas]

[El usuario selecciona una carrera profesional]

¡Genial! Ahora vamos a ver cuánta demanda tiene esta carrera

En [provincia usuario] hay [nº de ofertas] ofertas de trabajo para la carrera de [nombre carrera]

En España, el top 3 de provincias donde hay más demanda es:

[Gráfico con el top 3]

Ahora que ya sabes todo esto, ¿cómo quieres continuar?

[Botones: Quiero formarme, Ver en otras provincias, Ver otras carreras, Ir por destrezas]

OPCIÓN 1: Usuario escoge “Ver en otras provincias”

Dime la provincia en la que quieres ver la demanda de esta carrera

USUARIO: [provincia]

B: En [nombre provincia] hay [nº de ofertas] para la carrera [nombre carrera]

¿Quieres buscar en otra provincia?

[Botones: Sí, No]

OPCIÓN 1.1: Usuario escoge “sí”

B: Se le devuelve al primer mensaje de la opción 1

OPCIÓN 1.2: Usuario escoge “no”

B: ¡Vale! ¿Qué quieres hacer?

[Botones: Formarme, Ver otras carreras, Ir por destrezas y Otra cosa. El comportamiento es igual que los otros botones con el mismo nombre, si elige “otra cosa el agente le pregunta “¿qué quieres hacer?” y espera a la respuesta del usuario]

OPCIÓN 2: Usuario escoge “Ver otras carreras”

B: ¿Qué tipo de carreras quieres ver?

[Botones: Demandadas, Recomendadas, Ir por destrezas]

OPCIÓN 2.1: Usuario escoge “Demandadas”

B: Indícame si quieres verlas en España o en una provincia en concreto. En este caso, escribe el nombre de la provincia que quieres, si no, di “España”.

[Después de que el usuario diga la provincia o España, se saca hasta un máximo de 2 carruseles de 3 profesiones cada uno con los criterios que haya elegido el usuario]

OPCIÓN 2.2: Usuario escoge “Recomendadas”

[Se vuelve a enseñar el carrusel del principio y el flujo continúa]

OPCIÓN 2.3: Usuario escoge “Ir por destrezas”

Se va al caso de uso 5 – Destrezas

OPCIÓN 3: Usuario escoge “Ir por destrezas”

Se va al caso de uso 5 – Destrezas

OPCIÓN 4: Usuario escoge “Quiero formarme”

¡Has escogido una carrera profesional muy interesante!

Vamos a ver cuánta afinidad tienes con la carrera profesional que has seleccionado

Estas son las destrezas técnicas que requiere esta carrera:

- Destreza 1
- Destreza 2
- ...
- Destreza 10

Y estas son las que te faltan para perfeccionarla:

- Destreza 1
- Destreza 2
- ...
- Destreza 10

¿Sobre cuál quieres buscar cursos? Si no te interesa ninguna, escribe “ninguna”

[El usuario contesta con la destreza que quiere]

¡Genial! Ahora voy a mostrarte los mejores cursos sobre [nombre destreza]

[Se saca un carrusel de 3 cursos. Cada tarjeta debe tener el logo de la plataforma que ofrece el curso y un botón de “Ver detalle”, que abre la webview correspondiente].

¿Quieres ver más?

[Botones: Sí, No]

OPCIÓN 2: Usuario elige “No”

¿En qué más te puedo ayudar?

[Campo abierto, el usuario puede contestar con lo que quieras]

OPCIÓN 1: Usuario elige “Sí”

[Se saca un carrusel de 3 cursos. Cada tarjeta debe tener el logo de la plataforma que ofrece el curso y un botón de “Ver detalle”, que abre la webview correspondiente].

B: Si ninguno te interesa, puedes buscar más base a filtros o según destrezas. ¿Qué quieres hacer?

[Botones: Filtros, Destrezas]

OPCIÓN 1: Usuario elige “Destrezas”

B: Se va al caso de uso 5 – Destrezas

OPCIÓN 2: Usuario elige “Filtros”

Elige el filtro que quieres aplicar

[Botones: Precio, Modalidad y Nivel]

OPCIÓN 1: Usuario elige “Precio”

¿Cómo quieres que sean los cursos?

[Botones: Gratuito, Pago y Todos]

[Elija la opción que elija se saca hasta un máximo de dos carruseles con hasta 3 cursos cada uno con el filtro aplicado. El comportamiento es igual que con el carrusel normal de cursos del principio]

OPCIÓN 2: Usuario elige “Modalidad”

¿De qué tipo quieres que sean los cursos?

[Botones: Online, Presencial y Todos]

[Elija la opción que elija se saca hasta un máximo de dos carruseles con hasta 3 cursos cada uno con el filtro aplicado. El comportamiento es igual que con el carrusel normal de cursos del principio]

OPCIÓN 3: Usuario elige “Nivel”

¿Qué nivel quieres que tengan los cursos?

[Botones: Básico, Intermedio, Avanzado y Todos]

[Elija la opción que elija se saca hasta un máximo de dos carruseles con hasta 3 cursos cada uno con el filtro aplicado. El comportamiento es igual que con el carrusel normal de cursos del principio]

Anexo VI. Respuestas originales a la pregunta seis de la encuesta.

Para el caso del guión creo que sería mas cómodo contar con un guión por cada caso de uso o camino único que puede realizar el usuario, ya que en el formato actual es bastante complicado seguir los posibles flujos a implementar.

Para pasar a código los modelos necesitaría más especificación de los atributos de cada elemento

En el modelo del flujograma veo necesario añadir una leyenda que facilite al usuario el entendimiento del modelo sin necesidad de leerlo, solo viendo los coles.

En caso del Guión veo que es didifilmente comprensible porque:

- Faltaría mejorar formato para facilitar que se entienda en un vistazo si es el bot el usuario
- igualmente con deambiguaciones y comentarios del linguista, al no estar distinguidos cuesta enterder el flujo a la primera en una lectura rapida.

Incluiría los mensajes que introduce el usuario y una leyenda en los diagramas que indique qué es cada elemento del diagrama.

Flujograma:

- ¿Que continua tras registrarme? ¿Campo abierto? ¿Carreras?
- No se cómo se llega al flujo 2 "recomendación de carreras"
- No se cómo se llega al flujo 3 ya que empieza con "has escogido una carrera interesante", supongo que es tras el "quiero formarme"
- No se especifica que las carreras y los cursos deben llevar logos e imágenes ni que hacen los botones "ver video" ni "ver detalles" de cada uno (en el guion si se detalla esto)
- No se detalla con exactitud que ocurre tras seleccionar los filtros y mostrar el carrusel: ¿Campo abierto?¿Puedo escoger más filtros? ¿Lía me muestra algún botón? (En el guion si se detalla esto)

Guion:

- El orden de las opciones desorienta a veces y es fácil perderse
- Dividiría el guion según el flujo para entender de manera clara cuando estoy en el flujo de registro y cuando cambio al flujo de carreras por ejemplo

En caso de error que pasaría

Asignaría iconos(rombos, rectangulos, circulos) a cada interacción posible en cada punto, como una decisión, un inicio o final de conversación, etc

Las varias respuestas que pueden ocurrir tras una transición de estado a estado, deberían de estar representado en el diagrama de flujo, en vez de explicar que ha de verse un Excel.

Anexo VII. Respuestas originales a la pregunta ocho de la encuesta.

Desde el punto de vista del desarrollador, contar con el flujograma ayuda a poder interpretar correctamente todas las casuísticas a tener en cuenta, teniendo una visión global del proceso a implementar y quedando claros todos los casos de uso.

Es más claro, a primera vista se más rápido y es mucho más fácil de entender.

Me parece mucho más visual, fácil de entender de un vistazo, y mucho más cercano al lenguaje que utilizan los equipos de desarrollo para su posterior implementación.

Me parece mucho más fácil de entender visualmente.

Creo que está mejor representada la información y que es una interacción mucho más real con una persona. En cuanto a pasar a código quizá si que sea más difícil pero creo que tiene más ventajas en cuanto a usabilidad que el guión.

Es mucho más fácil de seguir, de visualizar y (sobretudo) aclara muchísimo más a la hora de programar. Gracias a ellos podría dividir mi código de manera eficaz en diferentes archivos de manera más ordenada y entendería que hace cada cosa.

El guion es muchísimo más caótico y es muy fácil perderse u olvidar que todavía quedaba una opción más que había aparecido anteriormente al listar los botones. Muchas veces también se pierde el hilo de si estamos en un flujo o si hemos salido de él, es decir, si las respuestas de Lía siguen siendo parte de la Opción X o si nos encontramos en un flujo a parte. De todas maneras escribir esto último de manera clara sería imposible ya que acabarían apareciendo cosas como "Opción 2.1.4.2".

El flujograma aun así necesita de más detalles en algunos puntos y creo que nunca se debería terminar un diagrama sin saber a donde va, es decir, los diagramas siempre deberían terminar en un "campo abierto", en un "Se va al caso de uso X" o en un "vuelve al punto X" (como hace el segundo diagrama con "quieres buscar otra provincia -> Si -> 2") ya que el programador no puede dejar el bot "muerto", siempre debe quedarse en un punto donde el usuario decide cual es la siguiente opción (ejemplo: cuando el usuario elige "formarme" el programador no sabe que debería hacer Lía). De todas maneras, esto ocurre en ocasiones puntuales y son cuestiones sencillas de resolver, pero para ello es necesario contactar con el/la lingüista encargado/a (es por ello que he marcado la opción afirmativa de necesitar ayuda al pasarlo a código).

En definitiva, el flujograma me ayudaría mucho más a la hora de programar a Lía y solo necesitaría ayuda externa para cuestiones y detalles más profundos.

El flujograma me parece un modelo mucho más visual y fácil de entender a la hora de desarrollar las funcionalidades que se plantean.

Porque tienes una visión global del flujo. Te permite diferenciar de un vistazo que partes se pueden desarrollar en paralelo y las que son parecidas para no implementar dos veces lo mismo. También te permite aislar zonas, es decir, al tener una vista en general del flujo puedes decidir de una manera rápida hasta donde puedes desarrollar o que partes hacer en el futuro sin dejar pobre la funcionalidad.

El flujograma contiene más información visual para el programador y da lugar a menos confusión que el guión

Queda más claro y de un vistazo puedes ver todo el caso de uso para poder tomar mejores decisiones técnicas. También añadiría más detalle para saber que debo mostrarle al usuario (texto, botones, imágenes, gif, etc) y además algún modo de poder ver la url/diseño de esos elementos gráficos.

Es una vista esquemática que se entiende muy claramente y su traducción a código es bastante directa.

Fácil comprensión y asimilación

Me parece una representación más limpia, que te permite crearte una idea clara en la cabeza de lo que debes conseguir hacer. Además te dibuja todas las opciones lógicas que posteriormente tendré que programar, mientras que en el guión tengo que leer varias veces, e ir buscando las opciones en el texto.